THE
PHYSICAL REVIEW
A Journal of Experimental and Theoretical Physics Established by E. L. Nichols in 1893
VOL. 52, NO. 3 AUGUST 1, 1937 SECOND SERIES

Ionization by Gamma-Rays and Cosmic Rays in Gases at High Pressure and High Collecting Fields

J. CLAY
Natuurkundig Laboratorium, Amsterdam
(Received April 8, 1937)

Experimental test and theoretical considerations show that for the gamma-ray ionization measurements at high pressures by Clay and van Tijn at collecting fields of 5000−400 volt/cm there was no volume recombination. These measurements cannot be affected by criticism of Bowen and Cox founded on experiments with collecting fields of 1009−1.5 volt/cm. The theory advanced by the author that for all pressures the volume ionization is proportional to the pressure and that the wall ionization (produced by electrons ejected from the wall) is for high pressures independent of the pressure is verified by experiments in air and in argon for gamma-rays and for cosmic rays. This result removes the doubt expressed by Bowen and Cox on this point. For obtaining the saturation values at high pressures the Jaffé theory provides the only reliable method in existence.

I

THE ionization currents in gases at high pressure have been the object of observation and discussion by several investigators.[1, 2] Such investigations became necessary, when high pressures were used for getting enough ions from radiations of very small intensity, such as the cosmic radiations.

For every future theory it will be of principal interest to know the real number of ions produced by the radiation, the saturation value of the ionization current. To this end one has to eliminate the volume recombination and the columnar recombination. Now it is easy to eliminate the volume recombination by the use of even moderate fields. It will be shown in the following that we must have had less than 1 percent volume recombination in air at 95 atmos. when we used gamma-rays producing 7×10^5 ions per cm^3 per sec. and applied a collecting field of 100 volt per cm. (See III.)

While for the volume recombination the intensity of the radiations is of principal importance, the case is quite different for the columnar recombination. This kind of recombination is independent of the intensity because in this case only the processes in each individual column play a role. For this reason the weakness of the radiation is of no help in obtaining saturation.

Thus both experiment and theory show that it is very difficult not to say impossible to avoid the influence of columnar recombination at high pressures, even when using high fields, excluded for argon, and other noble gases.

For instance, Jongen and I found that cosmic radiation[3] gives 40 ions per cc per sec. in air at 37 atmos. pressure under a shield of 12 cm Fe at sea level and 50° ML but that with a field of 200 volt per cm only 70 percent of the ions were drawn from the columns. Even at 600 volt/cm there still remains a columnar recombination of 20 percent. Volume recombination does not pro-

[1] J. Clay and M. v. Tijn, Physica **2**, 825 (1935).
[2] I. S. Bowen and E. F. Cox, Phys. Rev. **51**, 232 (1937).
[3] J. Clay and H. F. Jongen, Physica **4**, 245 (1937).

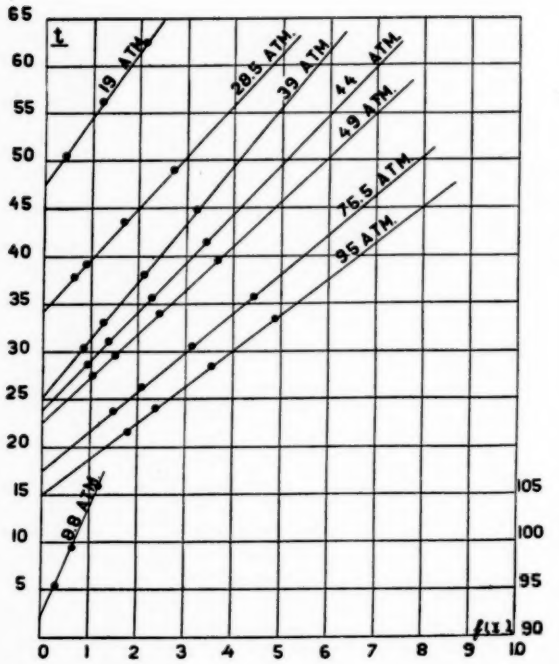


FIG. 1. Clay and v. Tijn. Measurements of ionization by gamma-rays in air between two flat plates at 0.6 cm distance at different pressures at different fields. According to formula 4 we may expect that the relation between I and $f(x)$ is linear for every pressure.

duce the slightest effect for such fields as we shall show further on.

II

It is of principal interest therefore to have a reliable way of finding the saturation value (that is, the ionization current produced by an infinite field), from the ionization found at finite fields. Now nearly every ionization (including that by gamma-rays and x-rays) takes place in columns as we see in Wilson photographs.

The theory of columnar ionization of Jaffé gives a good account of these phenomena, which have been proved in many cases. We think, therefore, the only way of reaching our aim is to follow this theory.

Zanstra,[4] using Erikson's measurements as an illustration, has found a very simple and useful way of so formulating Jaffé's theory, as to give the saturation currents. He found a linear relation between the inverse ionization and certain function of field and pressure, so that a linear extrapolation gives the saturation current I. The equation is

$$\frac{1}{i}=\frac{1}{I}+\frac{q}{I}f(x)(1)q=\frac{\alpha\cdot N_0}{8\pi D}. \tag{1a}$$

In this equation
α is the recombination coeff. in the gas,
D is the diffusion coeff. in the gas,
N_0 is the number of ions per cm in the column and

$$f(x)=e^{\frac{i\pi}{2}}H_0^{(1)}(ix)^{(2)}$$

where
$$x=\left(\frac{\text{b.u. sin }\varphi X}{2D}\right)^2. \tag{2a}$$

Further, u is the mobility of the ions, X the electric field, b the parameter of the column, and φ the angle between the axis of the column and X. As u and D are related to the pressure in the same way, and α is inversely proportional to p, Zanstra found it was possible to take for air

$$x=1,\ 24.10^{-4}\cdot(X/p)^2, \tag{3}$$

when X is expressed in volt/cm and p in atmospheres. As for $X=\infty$, $f(x)=0$, the value for the saturation current is found by taking $f(x)=0$ in (1).

When we measure the time t necessary to collect a given ionization charge and T represents the time for the same charge if we had saturation Eq. (1) assumes the form

$$t=T+qTf(x). \tag{4}$$

We have indeed shown in different series of experiments[1, 3, 5] that at a given pressure the values found are in the given linear proportion to $f(x)$ and in this way we have determined T (Figs. 1, 2 and 3). The linear relation (4) only holds for fairly homogeneous and fairly high fields. Needless to say the saturation value T can only be reliable when determined by an extrapolation from rather high fields where one is not too far from saturation.

III

We now come to the objection recently raised by Bowen and Cox[2] which is based on the work of the former. Bowen measured the ionization in air by gamma-rays but used a collecting field of

[4] H. Zanstra, Physica **2**, 817 (1935).

[5] J. Clay and G. v. Kleef, Physica **4**, Aug. (1937).

1009; 367; 230; 6.2 and 1.55 volt per cm. Only the field of 1009 volt in combination with higher fields would have been sufficient to test the formula of Jaffé, simplified by Zanstra and could be used to compare their values with ours which were at 5000, 3300, 1700 and 800 volt per cm.[6]

Bowen and Cox think that there was volume recombination in our case. There certainly was no appreciable volume recombination for the currents in the range between 800 and 5000 volt/cm which we used in our results, but also at the lower fields this recombination was much smaller than Bowen and Cox think. We can prove this theoretically and by experiment.

First we shall give an approximate theoretical argument. In the state of equilibrium let the number of ions present per cm³ be n, the velocity v, the area of the electrodes F. Then the ionization current will be

$$i = nFve.$$

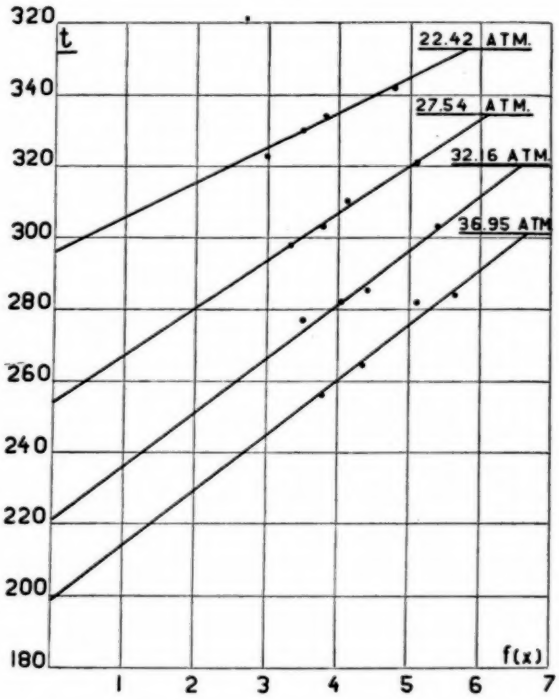


FIG. 2. Clay and Jongen. Measurement of ionization in air by cosmic rays between two coaxial cylinders at 3 cm distance at different pressures at different fields.

[6] It may be a question of taste that Bowen and Cox call the range 400–0 volt cm wide where the linear relation breaks down and is of no further use and that we are inclined to call the range infinity–400 volt/cm wide enough, where it holds and can be used successfully to find the saturation value.

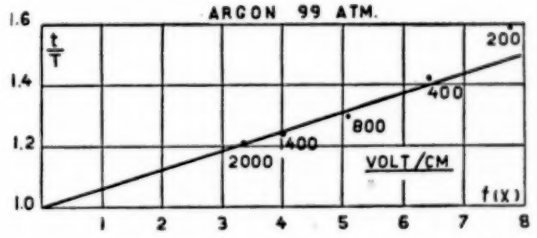


FIG. 3. Clay and v. Kleef. Measurements of ionization in argon by gamma-rays between two parallel grids at 0.5 cm distance. Between grids the deviation caused by volume recombination is readily noticed. Measurements are to be continued and published in *Physica* **4**, Aug. 1937.

When q pairs of ions are formed per cm³ per sec. the saturation current will be

$$I = qlFe,$$

where l is the distance between the electrodes. In our case the current i was a fraction a of I which was 0.33 at 95 atmos, and 166 volt/cm. From the above follows:

$$nFve = aqlFe,$$
$$n = aql/v$$

if u is the mobility of the ions at atmospheric pressure and V the potential

$$v = u_0 V/pl.$$

And when α is the recombination coefficient, which we suppose to have the same relation to pressure as u, we find for the fraction lost by volume recombination per sec.

$$\frac{\Delta n}{n} = \alpha n = \frac{\alpha_0}{p}\frac{aql^2p}{u_0 V}$$
$$= \frac{1.4\times10^{-6}\times0.33\times10^{6}\times0.36}{1.2V} = \frac{0.14}{V},$$

so that for $V = 100$ volt we would have a volume recombination of less than 1 percent.

Mr. v. Kleef and I tried also the volume recombination experimentally in the same way as was done by Bowen and Cox. We made two series of current measurements in air at 60 atmos. ionized by gamma-rays of 0.14 mg Ra at the distances of 26.7 cm and 10.2 cm.

The ionization was in the second case 7.2 times the ionization in the first case. We put the relation in the form

$$t/T = 1 + qf(x), \qquad q = \alpha N_0/4\pi d.$$

We see that for two sets of measurements the same relation will be found if the theory is right and the volume recombination is absent. This is indeed the case as is seen from Fig. 4, also at 200

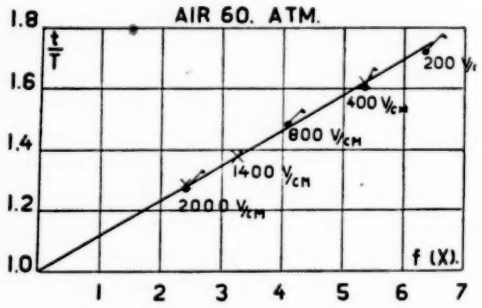


FIG. 4. Clay and v. Kleef. Measurements of ionization in air of 60 atmos. by two different gamma-radiations at different fields. ($\times I_1 = 7.9 \cdot 10^4$ e.s.u.; when possible the same sign as is given in the graph $\bullet I_2 = 5.7 \cdot 10^5$ e.s.u.).

volt/cm in the case of the strongest ionization ($5.7 \cdot 10^5$ I). Since, we have found the same evidence in argon of 100 atmos. In this case we had an ionization which was nearly 100 times as large, and we found a small volume recombination only for collecting fields smaller than 100 volt/cm.

We think we have now sufficiently replied to the first objections of Bowen and Cox since from a series of different experiments, as well as from theoretical arguments we have shown that we get the saturation values of ionization by the linear extrapolation method. Since the copy was sent to the *Physical Review*, Mr. v. Tijn and I found still another test by measuring Eve's constant, with a calibrated quantity of 23.16 mg Ra at 60 cm distance. From the saturation current found in air of 147.5 atmos. we found $4.30 \cdot 10^9$ I. (Physica **4**, Aug. (1937).)

IV

The second objection of Bowen and Cox is to the author's[7] theory of ionization as a function of pressure, the correctness of which they doubt.[8]

In this connection we can now discuss this theory of ionization which holds as far as our experiments show to pressures of 150 atmos. in air and in argon.

It was supposed by Bowen[9] and others that the ionization in a vessel should be directly proportional to the gas pressure only. But the experiments did not confirm this supposition, partly on account of lack of saturation.

Before this the opinion was put forward by Broxon[10] and others that the ionization was caused only by electrons projected from the atoms of the wall. In this case one has to expect that the ionization would be constant for pressures high enough, so that the whole energy of these electrons could be spent in ionization along their paths in the gas.

It was found by Broxon himself that this opinion could not explain the phenomena.

Then Bowen argued that if there were electrons that come from the wall, these electrons also would give an ionization proportional to the pressure so that the saturation ionization could only be proportional to the pressure but this opinion is not correct, as we will show on the basis of experiments.

The theory of ionization, proposed by the author,[7] takes into account the electrons of the gas and of the wall as principally different, so that the result was dependent on the pressure of the gas and the volume and the dimensions of the vessel.

The number of electrons ejected from the atoms of the gas will be proportional to the density of the gas.

The mean number of ion pairs given by one electron will be dependent on that part of the range of the electron that is enclosed in the ionization vessel, *viz.* on the dimensions of the vessel and the pressure of the gas. Let N_g be the average number of pairs that can be produced by one electron and p_0 the pressure which is necessary to limit the range in the vessel. Then the mean number of ion pairs for pressures below p_0 will be $N_g p/p_0$. If n_g is the number of electrons produced by the rays in one cc per sec. at one atmos., the total number of ions produced in the vessel of volume v, will be

$$I_g = A p^2/p_0 \quad \text{for} \quad p < p_0 \quad A = n_g N_g v. \quad (7a)$$

and $$I_g = A p \quad \text{for} \quad p > p_0.$$

[7] J. Clay, Physica **2**, 811 (1935).

[8] We are at a loss to follow their argument. Our theory does not presuppose that the efficiency of ejection of particles from the wall is larger than that from the gas, but rather that the range of the particles from the wall is much larger than of those from the gas.

[9] I. S. Bowen, Phys. Rev. **41**, 24 (1932).

[10] J. W. Broxon, Phys. Rev. **38**, 1704 (1931).

Now we consider the electrons from the wall. Let the number n_w of this kind be produced per cm² per sec. by the rays. If p_0' be the pressure for which the average range is just within the vessel and if the mean number of ion pairs for one electron is N_w, the mean number of pairs produced by one electron of the wall will be $N_w p/p_0'$.

The total number of pairs produced by the electrons from the wall, if we suppose that every part of the wall is on the same condition and its area F cm², is

$$I_w = Bp/p_0' \quad \text{for} \quad p < p_0' \quad B = Fn_w N_w.$$
$$\text{and} \quad I_w = B \quad \text{for} \quad p > p_0'.$$

In this simplified picture we have assumed that the wall particles ionize the gas in such a way that the ionization due to them is constant within a certain distance (the range) from the wall, and zero beyond this distance, not considering questions as to difference of directions and energy of the wall electrons.

The results of the experiments, where a flat condensor is used as ionization vessel, show that this assumption gives an approximation of the facts.

If the primary rays are photons, these will give pairs themselves. If the primary rays are charged particles which pass through the whole vessel as cosmic rays do, they will give a number of ions proportional to the pressure and proportional to the volume of the vessel. Adding the contributions (7), (8) and C_p we find for the total ionization

$$I = Ap^2/p_0 + Bp/p_0' + Cp \quad \text{for} \quad p < p_0,\ p_0',$$
$$I = Ap + Bp + Cp \quad \text{for} \quad p_0 < p < p_0',$$
$$I = Ap + B + Cp \quad \text{for} \quad p > p_0,\ p_0',$$

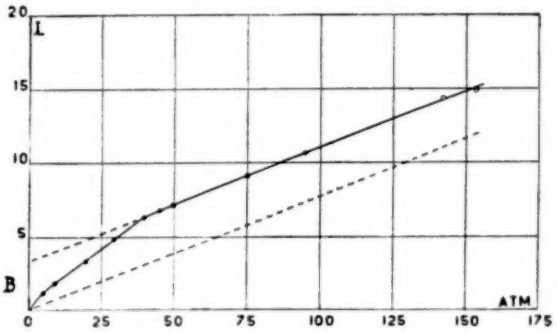


FIG. 5. Clay and v. Tijn. Measurements giving the relation between saturation current and pressure in air. Ionization by gamma-rays.

where A and B are given by (7a) and (8a) and C is zero if the primaries are photons. For high pressures we therefore expect that the relation of ionization to pressure is given by a straight line which does not pass through the origin, but intersects the I axis a distance B above it. This is

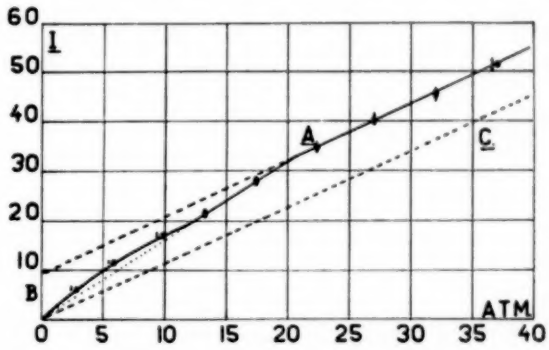


FIG. 6. Clay and Jongen. Measurements giving the relation between saturation current and pressure in air ionized by cosmic rays under a shield of 12.5 cm Fe.

proved by different experiments which we have already performed for gamma-rays in air (Fig. 5) and argon and for cosmic rays for argon and air.[11, 1, 3]

In order to find the ionization by cosmic rays in gas only the way we follow is the only way available at present. (Fig. 6.) This in reply to the criticism of Bowen and Cox who say that our method is not applicable to the problem of cosmic-ray ionization.

We have also realized experiments where the influence of the wall is eliminated by taking grid electrodes made of fixed wires. These experiments show a direct proportionality of the ionization to pressure, so that the factor B is negligible, as in this case it should be according to our theory (Fig. 8).

By taking grid electrodes and hydrogen we could also realize[7] the case for low pressures where

$$I = Ap^2/p_0 \quad \text{(formula (7) for } p < p_0\text{)}.$$

All of these relations are only satisfied when we measure the saturation current of ionization and we know that it is difficult to get saturation in most of the gases when the pressure is high. Only in argon can approximate saturation be reached

[11] J. Clay, Physica **2**, 111 (1935).

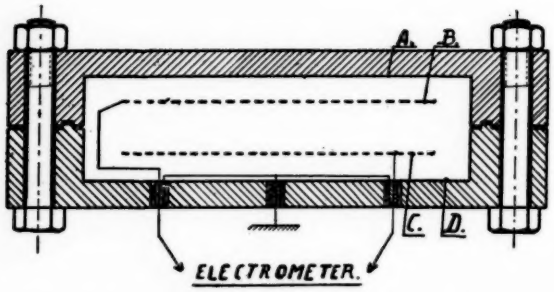


FIG. 7. Ionization vessel with two parallel grids.

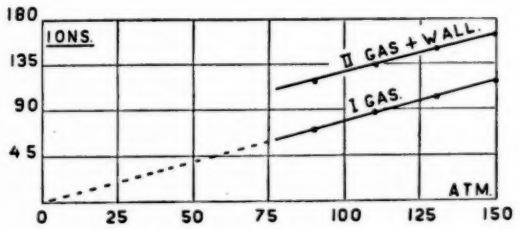


FIG. 8. Clay and de Bock. Saturation currents per cc in argon by cosmic rays under 16 cm Fe at different pressures. I. Between two grids. The ionization is proportional to pressure. II. The ionization between wall and grid $I=Ap+B$. The measurements are to be continued for the lower pressures and will be published in *Physica*.

at moderate fields. Therefore the criticism of Bowen and Cox about the linear relation

$$I=Ap+B$$

is shown to be groundless in the most direct way by experiments on ionization by cosmic rays in argon where we can find the approximate saturation values directly without the help of the Jaffé formula. This result was obtained first[11] in a vessel of the form given in Fig. 7. As long as the maximum range of the electrons is shorter than the distance, the ionization between the grids is directly proportional to the pressure, whereas the ionization between the wall and grid is increased by the ionization due to the electrons from the wall, which is constant above a pressure of 70 atmos.

Mr. de Bock and the author are now repeating this experiment with all possible precautions and have found that the ionization between the grids is accurately proportional to the pressure and that the ionization by the electrons between the wall and the grid is constant. (Fig. 8.) No objection can be raised to this result on account of eventual doubts regarding the Jaffé theory. Only we must have had a small lack of saturation which will not change this result. In our case we could not find a measurable difference between a collecting field of 200 and one of 650 volt/cm and we could not go higher.

V

It seems perhaps strange that there are such a small number of tracks of the electrons coming from the wall as seen on photographs of the Wilson chamber. To this we remark that the number of ion pairs produced by the electrons from the wall is for cosmic rays of the order of 100 per sec. per cm². When now the energy of one electron is 300,000 volt, one electron will make 10,000 ion pairs. We therefore need to have only one electron in 100 sec. per cm² to produce the ions. And, since a photograph takes about 1/25 sec., we have to look for one particle per 2500 cm² wall. Particles of this kind are certainly found. For particles from the wall ejected by gamma-rays we refer to the photographs of D. Skobelzyn.[12] He finds that these particles have generally a much greater energy than those ejected from the atoms of the gas. We hope to give in the near future a series of values of numbers of electrons ejected by the cosmic radiation from different materials.

[12] D. Skobelzyn, Zeits. f. Physik **28**, 278 (1924).

Scattering of Neutrons by Deuterons*

L. I. SCHIFF
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received June 7, 1937)

The elastic scattering of neutrons by deuterons is considered for energies small enough so that only the $l=0$ part of the incident wave is different from the corresponding part of a plane wave. The Breit-Feenberg Hamiltonian is used, assuming the same potential between all pairs of particles. Any possible polarization of the deuteron by the neutron is neglected, although the individual particles in the deuteron are taken into account by an approximate method. This method is capable of including exchange between the two neutrons. When the exchange term is included, the theory gives a cross section for thermal neutrons two to three times greater than observed; the cross section for 2.5 Mev neutrons is slightly greater than observed.

IN this paper we shall consider the elastic scattering of neutrons by deuterons at energies small enough so that only the $l=0$ part of the incident wave is different from the corresponding part of a plane wave. The interaction potential will be assumed to be the same for all pairs of particles[1] and to have the form used in previous researches;[2-4]

$$V(r)=J(r)[(1-g)P^M+gP^H]=J(r)[1-\tfrac{1}{2}g+2g\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2]P^M,$$
$$J(r)=De^{-2\epsilon r}-2De^{-\epsilon r},\quad \epsilon\equiv 2/r_0. \tag{1}$$

Here, P^M and P^H are the Majorana and Heisenberg operators, respectively, and the $\boldsymbol{\sigma}$'s are the Pauli spin matrices of amplitude $\frac{1}{2}$. Using nuclear units of length (8.97×10^{-13} cm) and of energy (0.506 Mev), we have $r_0=0.3$, $D=71.2$, and $g=0.3$.[5]

General Theory

Taking particle 1 to be the incident neutron and particle 3 to be the proton, we wish to find a solution of the wave equation (in nuclear units)

$$[-\tfrac{1}{2}(\Delta_1+\Delta_2+\Delta_3)+V(r_{12})+V(r_{13})+V(r_{23})-W_0-W]\Psi=0 \tag{2}$$

that has the asymptotic form (see Fig. 1a) as $r\to\infty$

$$\Psi\to[e^{ikr\cos\theta}+r^{-1}e^{ikr}f(\theta)]\phi_0(r_{23})S(123), \tag{3}$$

in the rest coordinate system (center of mass at rest). The beam of neutrons is directed along the negative z axis. S is a function involving only the spins (to be discussed in the next section), $\phi_0(r)$ is the wave function for the symmetric normal state of the deuteron with (negative) energy W_0, and W is the kinetic energy of the incident deuteron in the laboratory reference system (deuteron initially at rest).

* Preliminary report presented at the Washington meeting of the American Physical Society, April 28, 1937.
[1] Breit and Feenberg, Phys. Rev. **50**, 850 (1936).
[2] Morse, Fisk and Schiff, Phys. Rev. **50**, 748 (1936); **51**, 706 (1937). See also Ochiai, Phys. Rev. in press.
[3] Fisk, Schiff and Shockley, Phys. Rev. **50**, 1090 and 1191 (1936).
[4] Schiff, Phys. Rev. **51**, 783 (1937).
[5] The breadth used here ($r_0=2.7\times10^{-13}$ cm) is somewhat larger than the breadths used by other workers (Bethe and Bacher, Rev. Mod. Phys. **8**, 82 (1936), Feenberg and Share, Phys. Rev. **50**, 253 (1936), and Rarita and Present, Phys. Rev. **51**, 788 (1937)) even considering the difference in shape of the various potentials. The broader potential agrees (reference 2) with the older data (Harkins, Kamen, Newson, and Gans, Phys. Rev. **50**, 980 (1936), Kurie, Phys. Rev. **44**, 461 (1933)) on neutron-proton angle scattering. On the other hand, it seems to be incapable of explaining either the more recent data on this point (Bartlett, Phys. Rev. **51**, 889 (1937)), which is not yet certain, or the binding energy of H^3, if the variational method of Rarita and Present, Phys. Rev. **51**, 788 (1937) is as accurate as it appears to be. The calculations of the present paper were completed before the appearance of the paper of Rarita and Present. Since our calculations are essentially qualitative, it hardly seems worth while repeating them at this time. It does seem that the assumption of a narrower potential might bring the computed values closer to the experimental data on neutron-deuteron scattering.

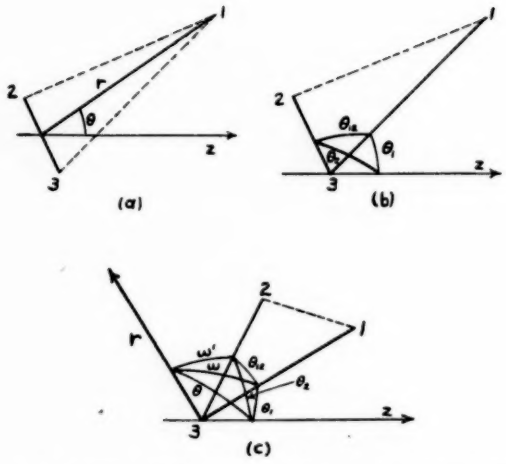


FIG. 1. The coordinate systems used in the theory. Particle 3 is the proton.

It soon becomes evident that in order to treat the problem in a reasonably simple manner, it is necessary to neglect any possible polarization of the deuteron by the incoming neutron, and to try to find a wave function which is the product of the normal deuteron function and a function for the scattered neutron. The coordinate system natural to the problem is that shown in Fig. 1a, in which the variables are r_{23} for the deuteron, and r and θ for the incoming neutron. The V's given by (1) involve space permutation operators, and will introduce functions of the medians of the (123) triangle other than r when they operate on a function of such coordinates. If a successive approximation method is used, this will make the resultant integrals exceedingly difficult to evaluate. This can be avoided by using the coordinate system of Fig. 1b, in which the incoming neutron is referred to the proton 3 rather than to the center of mass of the deuteron. The wave function then depends on the coordinates r_{23} of the deuteron and r_{13} and θ_1 of the incoming neutron; the space permutation operators introduce only the three sides of the (123) triangle into the successive approximation integrals. The Laplacian is no longer separable, but has the form

$$\tfrac{1}{2}(\Delta_1+\Delta_2+\Delta_3)\Psi(r_{23}, r_{13}, \theta_1)=\left\{\frac{1}{r_{23}^2}\frac{\partial}{\partial r_{23}}\left(r_{23}^2\frac{\partial}{\partial r_{23}}\right)+\frac{1}{r_{13}^2}\frac{\partial}{\partial r_{13}}\left(r_{13}^2\frac{\partial}{\partial r_{13}}\right)+\frac{1}{r_{13}^2\sin\theta_1}\frac{\partial}{\partial\theta_1}\left(\sin\theta_1\frac{\partial}{\partial\theta_1}\right)\right.$$
$$\left.+\cos\theta_{12}\frac{\partial^2}{\partial r_{13}\partial r_{23}}+\left(\frac{\cos\theta_2-\cos\theta_1\cos\theta_{12}}{r_{13}}\right)\frac{\partial^2}{\partial(\cos\theta_1)\partial r_{13}}\right\}\Psi(r_{23}, r_{13}, \theta_1). \quad (4)$$

Since we have decided to neglect polarization effects, (4) can be averaged over the orientation of the deuteron. Then the last two terms vanish, leaving

$$\tfrac{1}{2}(\Delta_1+\Delta_2+\Delta_3)\Psi(r_{23}, r_{13}, \theta_1)=(\Delta_{23}'+\Delta_1')\Psi(r_{23}, r_{13}, \theta_1), \quad (5)$$

where we adopt the definitions

$$\Delta_{23}'=\frac{1}{r_{23}^2}\frac{\partial}{\partial r_{23}}\left(r_{23}^2\frac{\partial}{\partial r_{23}}\right), \quad \Delta_1'=\frac{1}{r_{13}^2}\frac{\partial}{\partial r_{13}}\left(r_{13}^2\frac{\partial}{\partial r_{13}}\right)+\frac{1}{r_{13}^2\sin\theta_1}\frac{\partial}{\partial\theta_1}\left(\sin\theta_1\frac{\partial}{\partial\theta_1}\right). \quad (5')$$

Equation (3) gives the asymptotic form of the "ordinary" wave, in which the initially incident neutron is observed. There is also the "exchange" wave, in which the incident neutron changes places with the neutron initially bound in the deuteron, and the latter is observed. In the coordinate system of Fig. 1b, which we shall adopt throughout this paper, the asymptotic forms corresponding to the ordinary and exchange waves are

$$\Psi\rightarrow[e^{ikr_{13}\cos\theta_1}+r_{13}^{-1}e^{ikr_{13}}f(\theta_1)]\phi_0(r_{23})S(123), \quad \text{as } r_{13}\rightarrow\infty,$$
$$\text{and} \quad \Psi\rightarrow r_{23}^{-1}e^{ikr_{23}}g(\theta_2)\phi_0(r_{13})S'(123), \quad \text{as } r_{23}\rightarrow\infty. \quad (6)$$

When Ψ is made antisymmetric in the two neutrons (particles 1 and 2) as required by the Pauli principle, it has the asymptotic form when r_{13} (say) is large

$$\Psi(123)-\Psi(213)\rightarrow\phi_0(r_{23})\{e^{ikr_{12}\cos\theta_1}S(123)+r_{13}^{-1}e^{ikr_{13}}[f(\theta_1)S(123)-g(\theta_1)S'(213)]\}. \quad (7)$$

The forms (7) and (3) must of course involve the same k. On noting that the kinetic energy in the rest coordinate system is $2W/3$ and the reduced mass is $\frac{2}{3}$ of the neutron mass, it is readily seen that $k^2=8W/9$.

Since we are neglecting polarization and want Ψ to have the asymptotic forms (6), we write it as

$$\Psi=\phi_0(r_{23})F(r_{12},\theta_1)S(123)+\phi_0(r_{13})G(r_{23},\theta_2)S'(123)+\Phi, \quad (8)$$

where Φ involves a sum over excited and unbound states of the deuteron and will be neglected hereafter. We assume that in the zero-order approximation G vanishes and F is a solution of the equation

$$[\Delta_1'+k^2-U(r_{13})]F_0(r_{13},\theta_1)=0, \quad (9)$$

where U is the average field of the deuteron for the incoming neutron and can be written approximately as

$$U(r)=D_0e^{-2\epsilon_0 r}-2D_0e^{-\epsilon_0 r}. \quad (10)$$

It is clear that U must not contain a Majorana operator.

It has been shown by Morse[6] that the solution of (9) is given quite closely by a function of the form[7]

$$F_0(r,\theta)=(kr)^{-1}e^{i\delta}[\sin(kr+\delta)-e^{-\epsilon' r}\sin\delta\cos kr]+\sum_{l>0}(2l+1)i^lP_l(\cos\theta)j_l(kr) \quad (11)$$

whenever U has the general shape of a potential hole. For the potential (10), Morse[6] has found that when ϵ' is set equal to ϵ_0, and δ is obtained by setting $\int F_0(\Delta_1'+k^2-U)F_0d\tau$ equal to zero, the $l=0$ phase shift δ obtained in this way is never more than 7° from the phase shift obtained by exact solution of (9), and is generally less than 3° off, for a range of values of k, D_0 and ϵ_0. The depth D_0 of U depends strongly on spin orientation and on the manner in which the omission of the Majorana operator in U is corrected for; the mean breadth $(2/\epsilon_0)$ of U, however, depends principally on the mean breadth $(2/\epsilon)$ of the two-body potential (1). We take for the normal state deuteron function the normalized variational form[3]

$$\phi_0(r)=(a^3/\pi)^{\frac{1}{2}}e^{-ar},\quad a=3.236, \quad (12)$$

which gives within five percent of the deuteron binding energy (4.35 nuclear units) obtained from an exact solution of the deuteron equation

$$[\Delta_{23}'+W_0-V(r_{23})]\phi_0(r_{23})=0. \quad (12')$$

Then we can fix r_{13} and average the potential $[J(r_{12})+J(r_{13})]$ over r_{23} and θ_{12}. The resulting estimate $\epsilon_0=5.701$ is relatively insensitive to D and g. This ϵ_0 is somewhat smaller than the corresponding quantity $\epsilon=6.667$ appearing in (1), as would be expected.

Morse's method for finding the $l=0$ phase shift using the potential (10) would serve to give the scattering for energies up to the point where the $l=1$ phase shift becomes appreciable, if we had some knowledge of the depth D_0 of the equivalent field U which best accounts for the operator nature of the V's. D_0 can be estimated by writing (2) in a form which can be used to calculate the next approximation to F. Neglecting the G and Φ terms in (8), we obtain on substitution into (2), using (5) and (12′) and subtracting $U\phi_0FS$ from each side

$$[\Delta_1'+k^2-U(r_{13})]\phi_0(r_{23})F(r_{13},\theta_1)S(123)=[V(r_{12})+V(r_{13})-U(r_{13})]\phi_0(r_{23})F(r_{13},\theta_1)S(123). \quad (13)$$

[6] Morse, abstract 26 at the Washington meeting, April 28, 1937.

[7] Morse, *Vibration and Sound* (McGraw-Hill, 1936), p. 246, defines the functions: $j_m(z)=(\pi/2z)^{\frac{1}{2}}\cdot J_{m+\frac{1}{2}}(z)$, and gives several of their properties.

Multiplying both sides of (13) by $\phi_0^*(r_{23})S(123)$, integrating over r_{23}, and putting the approximate form F_0 for F on the right side, we obtain

$$[\Delta_1'+k^2-U(r_{13})]F_1(r_{13},\theta_1)$$
$$=\int\phi_0^*(r_{23})S(123)[V(r_{12})+V(r_{13})-U(r_{13})]\phi_0(r_{23})F_0(r_{13},\theta_1)S(123)d\tau_{23}. \quad (14)$$

Remembering that F_0 is a solution of (9), the asymptotic form of the solution of (14) is (see Fig. 1c)

$$F_1(r,\theta)\rightarrow F_0(r,\theta)-(4\pi r)^{-1}e^{ikr}\int F_0(r_{13},\pi-\omega)\phi_0^*(r_{23})S(123)[V(r_{12})+V(r_{13})-U(r_{13})]$$
$$\times\phi_0(r_{23})F_0(r_{13},\theta_1)S(123)d\tau_{123}, \quad \text{as } r\rightarrow\infty. \quad (15)$$

We now have an expression for the first approximation F_1 which involves the difference between the true potential $[V(r_{12})+V(r_{13})]$ and the average potential $U(r_{13})$. Then the best approximation to U from our point of view is that which makes the integral on the right side of (15) vanish, so that $F_1=F_0$.

A form like (15) for the first approximation to G can be obtained by similar methods. Neglecting the Φ term in (8), we obtain on substitution into (2), using the equations obtained from (5) and (12′) by interchange of 1 and 2

$$[\Delta_2'+k^2-V(r_{12})-V(r_{23})]\phi_0(r_{13})G(r_{13},\theta_2)S'(123)$$
$$=[V(r_{12})+V(r_{13})-\Delta_1'-k^2]\phi_0(r_{23})F(r_{13},\theta_1)\,S(123). \quad (16)$$

Putting approximately $V(r_{12})+V(r_{23})=U(r_{23})$ on the left side of (16), and F_0 for F and

$$(\Delta_1'+k^2)F_0(r_{13},\theta_1)=U(r_{13})F_0(r_{13},\theta_1)$$

on the right side, we obtain

$$[\Delta_2'+k^2-U(r_{23})]\phi_0(r_{13})G(r_{23},\theta_2)S'(123)=[V(r_{12})+V(r_{13})-U(r_{13})]\phi_0(r_{23})F_0(r_{13},\theta_1)S(123). \quad (17)$$

Multiplying both sides of (17) by $\phi_0^*(r_{13})S'(123)$ and integrating over r_{13}, we obtain an equation analogous to (14); the asymptotic form of its solution is (see Fig. 1c)

$$G_1(r,\theta)\rightarrow-(4\pi r)^{-1}e^{ikr}\int F_0(r_{23},\pi-\omega')\phi_0^*(r_{13})S'(123)[V(r_{12})+V(r_{13})-U(r_{13})]$$
$$\times\phi_0(r_{23})F_0(r_{13},\theta_1)S(123)d\tau_{123}, \quad \text{as } r\rightarrow\infty. \quad (18)$$

Spin Functions[8]

The spin function $S(123)$ must represent a neutron (particle 1) and a deuteron (particles 2 and 3) far apart, when the deuteron is known to be in a triplet state; it must also be an eigenfunction of the operator $\sigma_1\cdot(\sigma_2+\sigma_3)$ in order to remove off-diagonal matrix components which would prevent convergence of the successive approximations. The spin states divide into quartets (S_q) and doublets (S_d)

$$\begin{aligned}
&S_q^1=(+++),\\
&S_q^2=(1/\sqrt{3})[(++-)+(+-+)+(-++)],\\
&S_q^3=(1/\sqrt{3})[(--+)\\
&\qquad\qquad+(-+-)+(+--)],\\
&S_q^4=(---),\\
&S_d^1=(1/\sqrt{6})[(++-)+(+-+)-2(-++)],\\
&S_d^2=(1/\sqrt{6})[(--+)+(-+-)-2(+--)].
\end{aligned} \quad (19)$$

Here, a term such as $(+-+)$ signifies that spin state in which particles 1 and 3 have spin components whose eigenvalues are $+\frac{1}{2}$ along some arbitrary axis, and particle 2 has a spin component whose eigenvalue is $-\frac{1}{2}$ along the same axis. When neutron 2 is free, and particles

[8] The treatment given here is similar to that given in reference 4 in the section labeled "symmetry properties."

1 and 3 are bound in the deuteron (exchange wave), the spin states are given by (19) except that positions 1 and 2 are interchanged; thus $S'(123)=S(213)$.

If an energy M has the general form $A_0+B_0\sigma_1\cdot\sigma_2+C_0\sigma_1\cdot\sigma_3+D_0\sigma_2\cdot\sigma_3$, where A_0 ,etc., may be permuting functions of the space coordinates, it is readily shown that all matrix elements of this energy of the form SMS or $S'MS$ vanish unless the two spin functions correspond both to quartet states or both to doublet states. Thus only quartet-quartet and doublet-doublet transitions are possible. Since the final quartet and doublet states are orthogonal, we can compute the quartet and doublet scattering separately, and later combine the cross sections in the ratio 2 : 1 of their statistical weights. With the help of (7), we have the result that the whole differential cross section in the rest coordinate system is

$$\sigma(\theta)d\omega=\tfrac{1}{3}[2\,|f_q(\theta)-g_q(\theta)\,|^2+|f_d(\theta)-g_d(\theta)\,|^2]d\omega. \quad (20)$$

The subscripts refer to the use of quartet and doublet spin functions in calculating f from (15) and g from (18).

Results

The integrals (15) and (18) can be evaluated analytically for all energies (k values) with the help of the relation[3]

$$\int_v \frac{e^{-\alpha r_{12}-\beta r_{13}-\gamma r_{23}}}{r_{12}r_{13}r_{23}}d\tau_{123}=\frac{16\pi^2 v}{(\alpha+\beta)(\beta+\gamma)(\gamma+\alpha)}, \quad (21)$$

although the computations become exceedingly complicated for $k>0$. When $k=0$, (11) becomes

$$F_0(r,\theta)=1+(\Delta/r)(1-e^{-\epsilon' r}), \quad (22)$$

and the integrations are readily carried through. The f and g in (20) are then independent of θ, and the scattering is spherically symmetric in the rest coordinate system. In (22), $\Delta=\pm\lim(\sin\delta/k)$ as $k\to 0$, according as δ approaches an even or odd multiple of π. The criterion for determining whether $\delta\to 0, \pi, 2\pi, \cdots$ as $k\to 0$ is whether the potential hole U contains 0, 1, 2, $\cdots$ negative energy levels;[2] these levels may not correspond to physical reality. It turns out that the U's obtained below for both the quartet and doublet cases each contain one level, so that δ approaches π.

To determine D_0, we set $k=0$, $\epsilon'=\epsilon_0=5.701$, and find by trial and error a pair of values of D_0 and Δ that will make the integral on the right

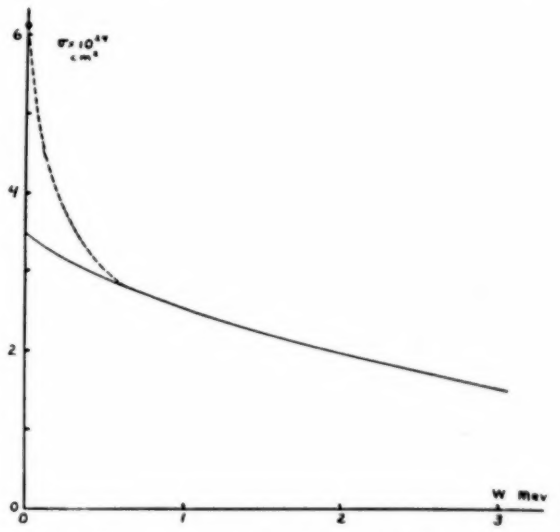


FIG. 2. Elastic scattering cross section of deuterons for neutrons as a function of incident neutron energy; without exchange (solid line), and with exchange (dotted line).

side of (15) zero and will also satisfy (9) according to Morse's criterion (vanishing of the average $H-k^2$). This must be done twice, using quartet and doublet spin functions. For the quartet case, $D_0=70.0$ and $\Delta=-0.564$; for the doublet case, $D_0=62.0$ and $\Delta=-0.629$. We can then say that the U's given by (10) with the above D_0's and $\epsilon_0=5.701$ in both cases are the best representations of the average field of a deuteron for an incoming neutron, when the coordinate system of Fig. 1b is used. For $k=0$, $f=\Delta$; for $k>0$, the $l=0$ phase shift can be found by Morse's method, in which case $f=\sin\delta/k$. The higher l phase shifts cannot be found by this method, so that our calculations are good only up to energies where the $l=1$ phase shift becomes appreciable. The total cross section computed in this way (neglecting exchange) is plotted as the solid line in Fig. 2. The effect of the exchange wave can be computed at zero energy by using (18) and (22). It is found that g is $+0.352$ for the quartet case, and -0.248 for the doublet case. The total elastic cross section at zero energy (including exchange) is, from (20), $(4\pi/3)[2\times(0.916)^2+(0.381)^2]=7.64$ square nuclear units or 6.11×10^{-24} cm^2. This is plotted as

the solid circle in Fig. 2. The evaluation of (18) is very difficult for $k>0$; however, it is reasonable to assume that the exchange effect drops off fairly rapidly with increasing energy. Although further calculations should be made on this point, the dotted curve shows roughly the expected behavior of the total cross section, including exchange. A further improvement in the theory could be effected by evaluating (15) as well as (18) for $k>0$ instead of using Morse's more approximate method; at the present time the increased accuracy does not seem to warrant the considerable computational labor that would be required.

It is interesting to note that the cross section at zero energy due to the initial quartet spin state is almost six times that due to the initial doublet spin state, even omitting the effect of the different statistical weights (see preceding paragraph). It seems likely that this arises principally from the short-range repulsion between neutrons with parallel spins due to the Pauli principle.

The only published experimental values for the neutron-deuteron collision are those of Dunning *et al.*,[9] which include capture as well as elastic scattering cross section. However there is good reason[10] to believe that the capture cross section is negligibly small compared to the elastic scattering cross section even at thermal energies. The experimental cross sections using D_2O are 4.0×10^{-24} cm^2 for thermal neutrons, and 1.6×10^{-24} cm^2 for neutrons of about 2.5 Mev energy. The effect of chemical binding of the deuteron in the D_2O molecule is negligible at the higher energy, but will serve to reduce the experimental value for thermal energy neutrons[11] by an uncertain factor that is probably between 1.5 and 2.0. Thus the theoretical value is two to three times too large at zero (thermal) energy, and about 25 percent too large at 2.5 Mev (see Fig. 2). The theory would give worse agreement than indicated at the higher energy if the $l=1$ phase shift were appreciable, but it seems unlikely that it will be large at this energy.

In conclusion, we can say that the agreement between theory and experiment is at least qualitatively satisfactory, and that it seems likely that a more accurate theory (taking polarization into account, for example) will be capable of explaining the experiments quantitatively. The assumption of a narrower and deeper potential between pairs of particles[5] would probably make the cross sections calculated here somewhat smaller, improving the agreement with experiment.

It is a pleasure to thank Professor Philip M. Morse for helpful discussions of many points of the theory, particularly in connection with the positive energy wave function.

[9] Dunning, Pegram, Fink and Mitchell, Phys. Rev. **48**, 265 (1935). I am indebted to Professor Dunning for confirming these values recently.

[10] Kikuchi, Aoki and Takeda, Tokyo Inst. Phys. and Chem. Research **31**, 195 (1937).

[11] Fermi, Ricerca Scient. **2**, 13 (1936). The ratio 4 : 1 given by Fermi between the cross sections of slow neutrons on rigidly bound protons and on free protons, becomes 2.25 : 1 for the neutron-deuteron collision. The precise value of the ratio of the observed cross section to that for a free deuteron depends on the molecular energy levels and on the neutron energy; it is probably between 1.5 and 2.0

On the Ultraviolet Light Theory of Magnetic Storms

A. G. McNish
Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, D. C.
(Received March 25, 1936)

The three notable solar flares observed in the western hemisphere in 1936 did not produce magnetic storms. It is inferred from terrestrial effects accompanying them that intense ultraviolet emission occurred at the same time. Study of the effects suggests that, when a flare of ultraviolet light increases the conductivity of the lower regions of the atmosphere, increased diurnal-variation currents flow. The current-systems necessary to produce magnetic storms are of an entirely different type—hence the observed effects do not support the ultraviolet light theory of magnetic storms.

IN a series of papers which have appeared in this journal E. O. Hulburt and H. B. Maris have proposed and developed a promising theory, attributing auroral phenomena and magnetic storms to flares of ultraviolet light from the sun, as a substitute for the earlier corpuscular theory which has encountered numerous difficulties. The theory, as first proposed by Hulburt,[1] suggested that aurorae are due to ions, formed in the upper fringe of the ionosphere in low and middle latitudes, which drift along the lines of magnetic force to lesser heights in polar regions and there excite the auroral light by collision with other molecules. Chapman,[2] in criticism of this preliminary paper, pointed out that the lines of magnetic force along which ions must drift to fall into polar regions pass many tens of thousands of kilometers above the earth's surface in equatorial and middle-latitude regions, at which heights the atmospheric density must be quite inappreciable.

Subsequently Maris and Hulburt[3] and Hulburt[4] met these objections and extended the theory to account for magnetic storms as well as aurorae, introducing the assumption that through collisions of the second kind, involving excited atoms, some of the atmospheric atoms are projected to the necessary heights. The number of atoms above the 450-km level was estimated as 10^{16} in a column one square centimeter in cross section. In this region 2×10^{6} atoms per second were assumed to acquire velocities as great as 10 km per second. A sudden flare of ultraviolet light from the sun, persisting for about half an hour, was assumed to cause an increase in the ionization and in the number of high flying atoms, the number of long-free-path ions thus produced reaching 10^{16} in a unit column. These ions, through the crossed gravitational and magnetic fields in equatorial regions, would give rise to an eastward electric current producing the increase in the earth's horizontal magnetic field observed at the beginning of a magnetic storm, while electric currents induced in the earth would cause the greater decrease in the field observed during the main phase. Falling from their great heights into polar regions, the high flying ions would cause aurorae and produce certain magnetic effects through their diamagnetic action.

Chapman[5] in further criticism of the theory stated that the explanation offered for the second-phase variations of a magnetic storm was contrary to electromagnetic theory and called attention to the fact that the curves of magnetic variation in polar regions, which the theory had explained, were incorrectly drawn by the authors. Dissatisfaction with the original hypotheses was also expressed but not extensively discussed since the theory in the form offered appeared incapable of accounting for the observed facts.

Without altering the original hypotheses Hulburt[6] in another paper described other effects involved so that the magnetic variations produced in accord with the more highly developed aspect of the theory agreed closely with the

[1] E. O. Hulburt, Phys. Rev. **31**, 1038 (1928).
[2] S. Chapman, Phys. Rev. **32**, 993 (1928).
[3] H. B. Maris and E. O. Hulburt, Phys. Rev. **33**, 412 (1929).
[4] E. O. Hulburt, Phys. Rev. **34**, 344 (1929).
[5] S. Chapman, Mon. Not. R. Astr. Soc., Geophys. Sup. **2**, 296 (1930).
[6] E. O. Hulburt, Phys. Rev. **36**, 1560 (1930).

observed facts, rejecting the first assumption of diamagnetic action in polar regions which did not give variations of the correct sign.

Since all ideas concerning the extreme upper ionosphere are highly speculative, a satisfactory appraisal of the theory and the underlying hypotheses on theoretical grounds is difficult. Direct observation of the terrestrial-magnetic effects produced by a flare of ultraviolet light from the sun should furnish a suitable basis for judging the theory. During 1936 three notable occurrences of *H*-alpha brightening in the region of sunspots were observed in the western hemisphere, all of which were attended by marked magnetic effects and fade-outs of radio signals from the ionosphere. These *H*-alpha brightenings are considered the visible manifestations of solar disturbances which emit intense flares of ultraviolet light, the latter, through its ionizing effect, being responsible for the changes in the earth's magnetism and the radio phenomena.[7]

These solar flares seem to be exactly what are called for by the ultraviolet theory of magnetic storms. A detailed study of the magnetic effects associated with them clearly indicates that *they do not produce magnetic storms but result in a very special type of magnetic effect* different in nature from those occurring during magnetic storms. In a recent review of theories of magnetic variations and aurorae Hulburt[8] states that a world-wide magnetic storm began simultaneously with the notable outburst on the sun occurring at $16^h\ 46^m$ GMT on April 8, 1936. He further concludes "the simultaneity of incipience of radio fade-out, ionospheric change, magnetic disturbance, and solar eruption appear to offer detailed support of the ultraviolet theory," adding the reservation: "However, rather than to stress the agreement prematurely it is better to await the results of future experiment." On the contrary, detailed reports by observers state "the magnetic traces for the preceding and following days, April 7 and 8, Greenwich dates, were quite undisturbed"[9] and "no magnetic disturbance followed."[10] Most magnetic storms persist for a day or more.

Evidence that a world-wide magnetic storm did not begin at the time stated is supplied by Fig. 1, showing the magnetic traces at the time of the solar eruption at three widely separated observatories, Huancayo (Peru), Cheltenham (United States), and Watheroo (Australia), and Fig. 2, showing the minimum degree of disturbance which is called a magnetic storm at those stations. This demonstrates that no storm accompanied or followed the solar flare. The curves in Fig. 1, except for the feature immediately associated in time with the solar flare, are

[7] J. A. Fleming, Terr. Mag. **41**, 404 (1936); A. G. McNish, Nature **139**, 244 (1937). Attention was first called to these phenomena by J. H. Dellinger, Phys. Rev. **48**, 705 (1935).

[8] E. O. Hulburt, Rev. Mod. Phys. **9**, 44 (1937).

[9] O. W. Torreson, W. E. Scott, and H. E. Stanton, Terr. Mag. **41**, 199 (1936).

[10] R. S. Richardson, Terr. Mag. **41**, 197 (1936).

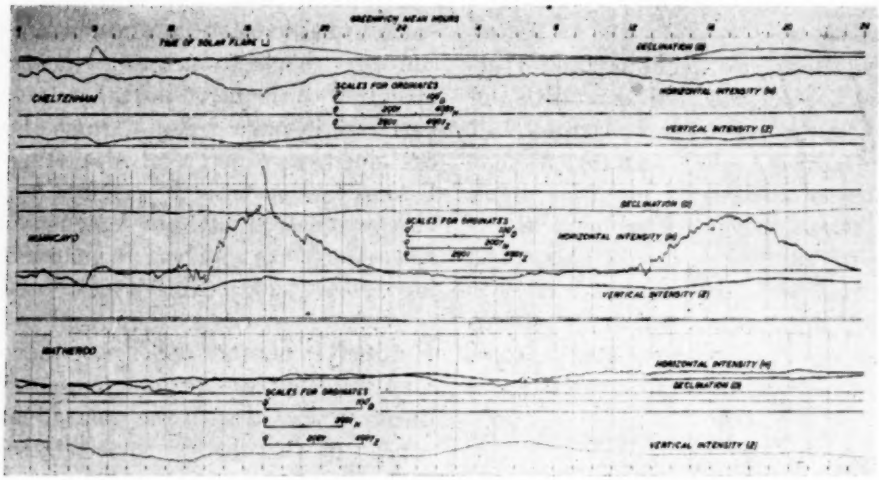


FIG. 1. Variations in earth's magnetic field before and after appearance of solar flare from $16^h\ 45^m$ to $17^h\ 03^m$ GMT, April 8, 1936.

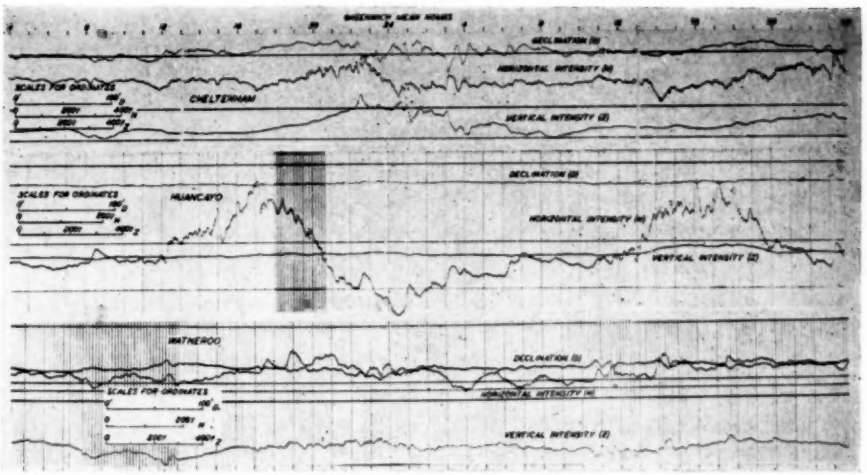


FIG. 2. Variations in earth's magnetic field during mild magnetic storm.

characteristic of what is observed at each observatory on an ordinary magnetically quiet day. No vestige of this solar-flare disturbance is present in the Watheroo record, indicating that the effect is not world-wide.

Careful examination of the records in Fig. 1 discloses that the earth's field was fluctuating slightly up to the time of occurrence of the flare but after its occurrence the earth's magnetic condition was even more quiescent. The records from 12 other magnetic observatories, scattered over the earth, have been examined. They reveal the same conditions. The pronounced but short-lived jump in the records, coincident with the flare, cannot be regarded as a magnetic storm, in the ordinary sense in which the term is used; nor is it world-wide, being confined to a region less than 90° from the subsolar point at time of occurrence. (Note absence of effect at Watheroo.)

Thus the statement in Hulburt's paper that a world-wide magnetic storm began at the time of fade-out is contradicted by the observed facts.

Plots of the average intensity of magnetic disturbance for each day at the Cheltenham, Honolulu, Huancayo, San Juan, Sitka, Tucson, and Watheroo magnetic observatories during the months of April, August, and November 1936 (Fig. 3) show that none of the three notable solar flares of the past year occurred during or immediately preceding a storm. Each occurred during a period of comparative quiescence and was immediately followed by fairly quiet days. Since all of these were exceptionally strong flares, one would have expected the magnetic disturbance to attain storm levels for the next two or three days in accord with the ultraviolet light theory.

The magnetic effects associated with the solar flares are found to be of a special type, distinct from the types accompanying magnetic storms. They consist of changes in the magnetic force at each station closely simulating in direction and magnitude the departures obtaining at the time of occurrence due to the regular diurnal variations of terrestrial magnetism which are strongly present even during the most quiet magnetic conditions. These ordinary diurnal variations, according to the Stewart-Schuster theory, are

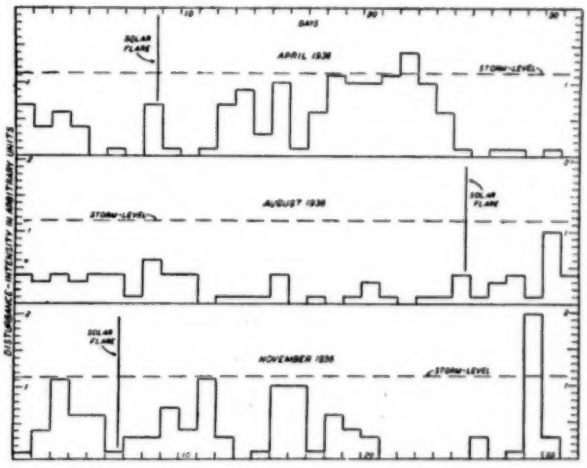


FIG. 3. Intensity magnetic disturbance, April, August, and November, 1936. based on numerical ratings from seven American-operated observatories.

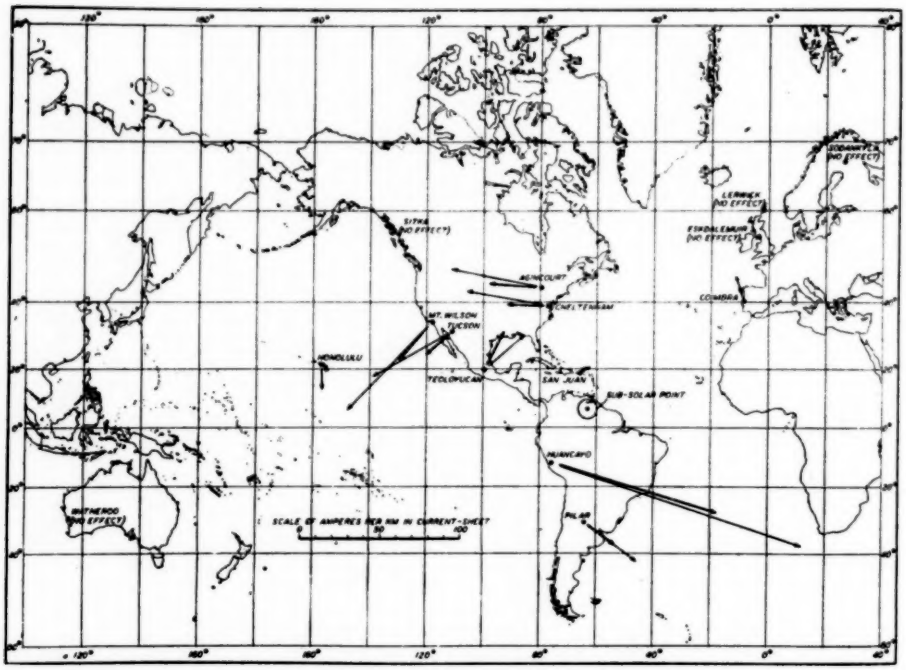


FIG. 4. Overhead currents necessary to produce magnetic changes observed at time of solar flare beginning $16^h\ 46^m$ GMT, April 8, 1936 (heavy arrows), and normal diurnal-variation changes at same time (light arrows).

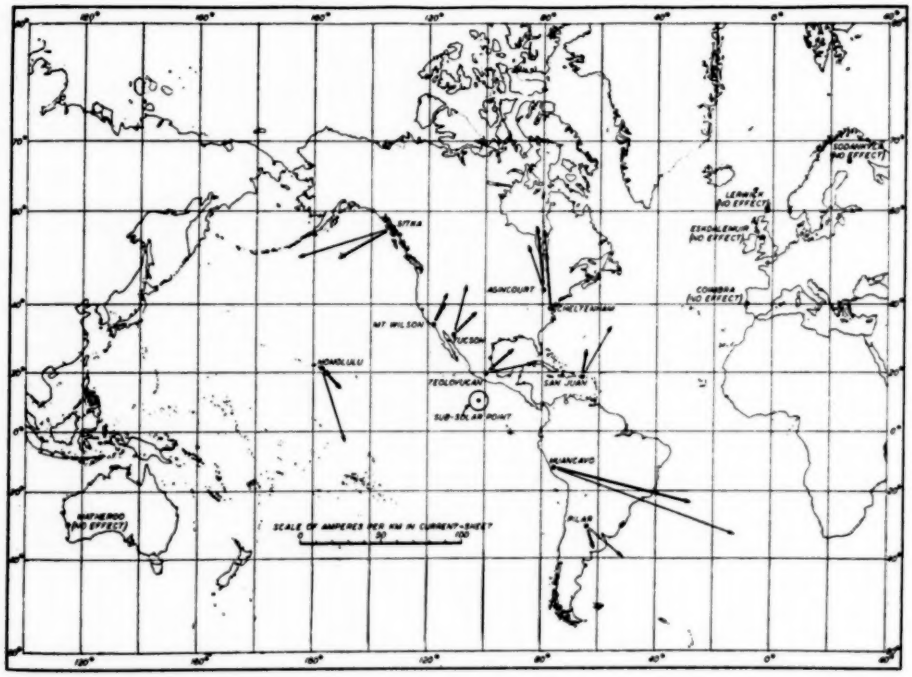


FIG. 5. Overhead currents necessary to produce magnetic changes observed at time of solar flare beginning $18^h\ 25^m$ GMT, August 25, 1936 (heavy arrows), and normal diurnal-variation changes at same time (light arrows).

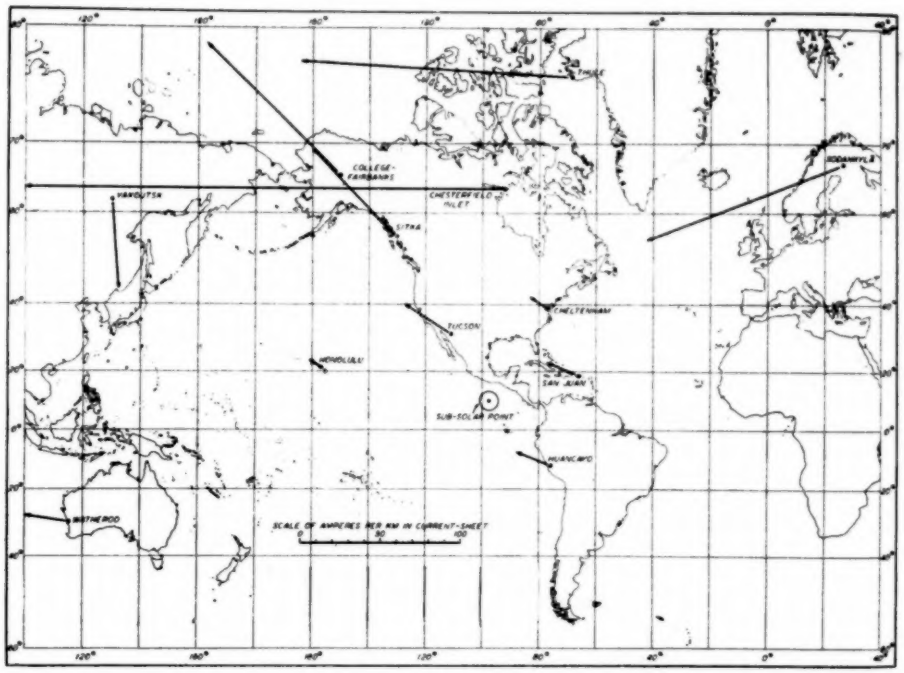


FIG. 6. Overhead currents necessary to produce magnetic changes observed at time of "bay"-type disturbance at $19^h\ 35^m$ GMT, April 16, 1933 (this type frequently occurs during magnetic storms).

caused by currents flowing in the atmosphere rendered conducting by solar ultraviolet light. A sudden increase in the ionizing radiation from the sun would cause an increase in these currents through the improved conductivity of the upper atmosphere. This view is supported by Figs. 4 and 5, showing the overhead currents necessary to produce the field changes associated with solar-flare disturbances (heavy arrows) and those necessary to produce the normal diurnal variation departures obtaining at the time (light arrows).[11] This increase of ionization occurs in the short-free-path region of the ionosphere and thus causes absorption of radio waves. It accounts for the observed fact that radio waves at frequencies which would ordinarily penetrate that region and be reflected back to earth from one or the other of the higher layers of ionization are not reflected during the disturbed period. Later, when the phenomenon has subsided, the return of reflections from those upper regions reveals that these layers have experienced no noticeable change.

On the other hand, the field changes occurring during a magnetic storm must involve a different mechanism. The overhead currents necessary to

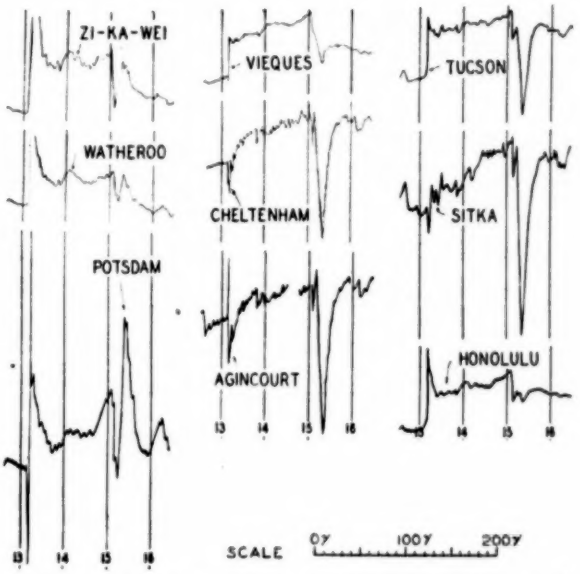


FIG. 7. World-wide sudden commencement of magnetic storm at $13^h\ 09^m$ GMT, May 13, 1921, horizontal-intensity changes.

[11] A detailed study of the magnetic effects produced by these solar flares is under way at the Department of Terrestrial Magnetism, the results of which will be published shortly in *Terrestrial Magnetism and Atmospheric Electricity*.

produce the "bay"-type magnetic disturbance which frequently occurs during magnetic storms are shown in Fig. 6.[12] While the currents necessary to produce the solar-flare disturbance are greatest close to the subsolar point and do not flow beyond the twilight zone, those necessary to produce the bay-type disturbance are much more pronounced near the auroral zone (about 70° north) and extend over the entire earth. The form of the apparent system of circulation is obviously different in both cases.

Actual records of the sudden commencement of a magnetic storm at various observatories are shown in Fig. 7. This represents another type of magnetic disturbance quite different from the other two, but one which is characteristic of the beginning of a magnetic storm. The outstanding characteristic of this disturbance is its widespread manifestation, in sharp contrast to the circumscribed manifestation of a solar-flare disturbance. Obviously, the mechanism which gives rise to the sudden commencement of a magnetic storm must be quite different from that which operates when a sudden flare of ultraviolet light from the sun impinges on the earth's atmosphere.

Acknowledgment.—The writer expresses his appreciation to various individuals and organizations which have made available to him the special data necessary for this study, and in particular to Dr. J. A. Fleming, Director of Carnegie Institution's Department of Terrestrial Magnetism, whose recognition of the importance of these phenomena has caused him to encourage and support the investigation of which this is a part.

[12] A. G. McNish, Trans. Am. Geophys. Union, 17th annual meeting, 166 (1936).

AUGUST 1, 1937 PHYSICAL REVIEW VOLUME 52

The Far Infrared Spectrum of Water Vapor

H. M. RANDALL, D. M. DENNISON, NATHAN GINSBURG AND LOUIS R. WEBER
University of Michigan, Ann Arbor, Michigan

(Received May 8, 1937)

The rotation spectrum of water vapor has been measured with high dispersion from 18μ to 75μ. The spectrometer employed echelette gratings and higher spectral orders were successfully removed through the use of filters and selective reflection. The positions of single lines were determined with an accuracy of about 0.05 cm^{-1} throughout the entire region while neighboring lines as close together as 0.5 cm^{-1} could be separated and measured. The analysis of the spectrum was accomplished in the following manner. From the values of the moments of inertia found by Mecke the energy levels were calculated from the asymmetric rotator equations. However, these computed levels are often far from the actual energy levels due to the presence of a large correction arising out of the centrifugal force stretching of the molecule. An estimate of this correction, which amounted in some cases to over 200 cm^{-1}, was made, thus furnishing energy levels with which to begin the analysis. The intensities of the rotation lines were calculated by using as an approximation the symmetric rotator amplitudes. By comparing the expected spectrum with the observed spectrum it was then possible to identify the lines and through them to determine the actual energy levels of the water vapor molecule. Combination relations as well as the formation of analytic series from analogous lines, served as important checks on the identifications. Through these methods the rotational energy levels of the water molecule have been found up to and including the group of $J=11$ with an accuracy around 0.1 cm^{-1}. The highest level determined, possesses an energy of over 3200 cm^{-1}, while all the rotational levels with energies less than 2000 cm^{-1} have been obtained. Finally all the allowed transitions together with their intensities are calculated and these are plotted directly above the observed spectrum. The agreement is remarkably good. All the essential features of the spectrum and indeed most of the finer details are correctly reproduced.

I. EXPERIMENTAL

THE present investigation covers the spectral region extending from 18μ to 75μ, thus overlapping somewhat the work of Wright and Randall[1] which began at 60μ. This earlier work demonstrated that the equipment used was capable of resolving and measuring the rotation spectrum of water with a completeness and accuracy not hitherto attained in this region of the spectrum. It was, however, so responsive to outside disturbances that much time was lost waiting for favorable conditions, often to be

[1] Norman Wright and H. M. Randall, Phys. Rev. **44**, 39 (1933).

obtained at night only. The spectrograph employed was the one designed by Randall[2] and following this first investigation he was able to find means of overcoming this defect of unsteadiness so that observations are now generally possible at any time during the day: only very seldom are outside disturbances of sufficient magnitude as to interrupt work. In Fig. 1 the first set of data taken in this investigation is shown. The grating was turned by steps of 1 minute of arc and a *single* galvanometer deflection only was taken for each position. The definiteness with which each line is determined clearly indicates the steadiness of the instrument. The seven lines lie between 22.5 and 23.6μ, covering a frequency range of 20 cm^{-1}. This very complete freedom from chance effects has permitted an amplification of from 150 to 300 times, which is somewhat greater than was previously used. This makes narrower slits possible with a corresponding increase in resolution. The measurable limit of 1 cm^{-1} between the lines of a narrow doublet established in the investigation of Wright and Randall is now reduced to 0.5 cm^{-1}.

By using essentially the same technique as in the earlier work the accuracy now attained in measuring sharp lines which are free from overlapping by adjacent lines is about 0.003μ in the shorter wave-length region, with less accuracy as 70μ is approached. The accuracy in terms of waves per cm appears to be about 0.05 cm^{-1} throughout the spectrum measured. The slit width of 0.5 mm at the center of the region measured with the 1200 line grating, namely at about 23μ, covered a spectral range of approximately 0.005μ. At 34μ, using the 900 line grating, the slit was 1 mm wide and corresponded to a spectral aperture of 0.01μ. Between 40 and 75μ the slits were 1.5 mm wide which with the 360 line grating meant a spectral range of 0.05μ. Under these circumstances, lines 1.5 cm^{-1} apart are completely resolved, while those but 0.5 cm^{-1} apart are clearly double and quite accurately measurable.

The 1200 line grating, having a ruled surface of 9×9 inches, was designed to locate the blaze or region of maximum energy at 18μ. The 900 line grating, surface 10×20 inches, had its blaze at 23.4μ while the blaze of the 360 line grating, surface 10×20 inches, was located at 57μ. In general it is advisable to work to one side or the other of the blaze, as the radiation of short wave-length concentrated at this place is so great that the devices adopted, to purify the spectrum elsewhere, are not always effective here. The methods of calibration were those described in the first paper.

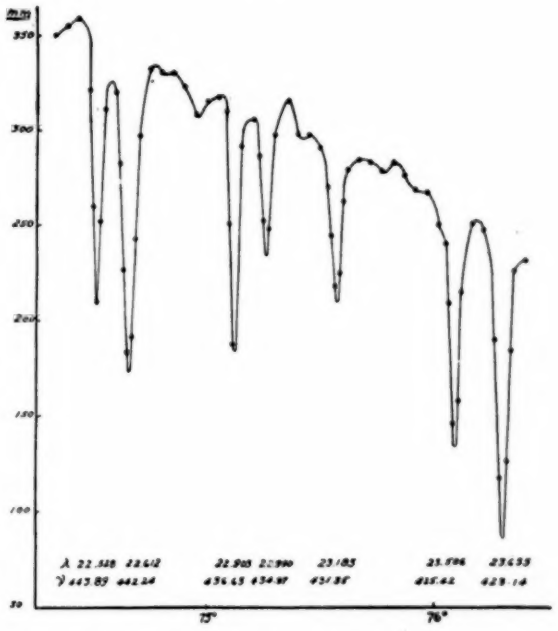


FIG. 1. The rotation spectrum of water.

Echelette gratings may be ruled to conserve the radiation in any spectral region but the desired first-order spectrum is overlaid with higher orders of radiation of shorter wave-length. These higher orders are removed in large part by absorbing screens and by selective reflections. The remaining energy of the higher orders is neutralized by employing shutters transparent to them but opaque to the longer wave-length radiation. Since this procedure is generally wasteful of the desired first-order spectra, the relatively large quantity of energy made available by this spectrometer through its large aperture and gratings permits not only greater dispersion but also greater losses in order to obtain the required degree of purity of spectrum. In the region between 18 and 30μ, a paraffin filter 1 mm thick, selective reflection from fluorite and a shutter of rocksalt cut out all

[2] H. M. Randall, Rev. Sci. Inst. **3**, 196 (1932).

observable effects of higher orders. Moreover the absorption of water vapor in the region between 9 and 15μ is so slight that second-order lines of measurable intensity are not likely to occur. To work between 25 and 40μ the only change necessary is that the selective reflection from a plate of sodium fluorite replaced that from fluorite. From 40 to 70μ, filters of paraffin and of crystal quartz each 1 mm thick together with shutters, first of potassium chloride, and then of potassium bromide, gave very satisfactory results. With a spectrum so rich in lines it is inevitable that the higher orders of some of the shorter wave-length lines will have values which will be in close agreement with the measured values of first order lines of longer wave-length. A careful study has been made of all such cases and it is believed that the results given have not been measurably affected by such possible overlappings.

This investigation was begun in April, 1934, and continued through the early summer. A sabbatical leave interrupted the work and in the meantime, June 15, 1935, a paper in this same field by Barnes, Benedict and Lewis[3] was published. From this paper it seemed quite probable that the higher resolution given by the spectrograph here used would be necessary if a successful interpretation of the rotation spectrum of water vapor in the far infrared was to be made. The work was accordingly continued.

II. Analysis of the Water Spectrum

In the remaining portion of this paper we shall attempt to identify the lines which have been found experimentally and thereby to determine the rotational energy levels of the water molecule. The measurements obtained in the present investigation will be combined with results found by Wright and Randall. The region thus included extends from 74 cm^{-1} to 560 cm^{-1} and contains some 173 lines. Since the average spacing between the lines is only 2.7 cm^{-1}, it is evident that the success and completeness of an analysis depends upon experimental data of high accuracy and resolution such as are now for the first time available.

1. Calculation of the energy levels

The starting point of this work rests upon the investigation of the vibration rotation bands of water vapor made by Mecke[4] and his collaborators. They have determined 42 of the rotational energy levels of the nonvibrating molecule including the states from $J=0$ to $J=4$ together with a few of the higher levels. A preliminary study showed that the rotation lines predicted from their levels were in most cases in good agreement with the positions of certain of the lines observed in the far infrared spectrum. It was therefore concluded that the moments of inertia deduced by them might be taken as substantially correct. This is the essential key without which the analysis would be impossible, for, through the moments of inertia the energy levels of the asymmetrical rotator may be calculated. The calculated levels then form the guides in determining the actual energy levels whose differences will give the rotation lines.

It was found convenient to employ the equations obtained by Nielsen[5] from the Wang secular determinant. In the course of the work a number of errors in Nielsen's equations were located. Professor Nielsen has subsequently informed us that these errors were typographical in nature. The list of corrections which he kindly sent us agrees with ours and is as follows.

$J=6$

2nd equation

$+4725b$ in the constant coefficient should be $-4725b$.

4th equation

$-636b^2$ in the coefficient of W^2 should be $-1176b^2$

$-34144b^2$ in the coefficient of W should be $-53664b^2$.

$J=9$

1st equation

-147456 in the constant coefficient should be $+147456$.

2nd equation

$+7056b^2$ in the coefficient of W^3 should be $-7056b^2$.

3rd equation

$+822149b^4$ in the coefficient of W should be $+1822149b^4$

$-45052805b^4$ in the constant coefficient should be $-115052805b^4$.

4th equation

$+822149b^4$ in the coefficient of W should be $+1822149b^4$

$-189344780b^2$ in the constant coefficient should be $+189344790b^2$

$-45052805b^4$ in the constant coefficient should be $-115052805b^4$.

[3] R. B. Barnes, W. S. Benedict, and C. M. Lewis, Phys. Rev. **47**, 918 (1935).

[4] R. Mecke, W. Baumann, and K. Freudenberg, Zeits. f. Physik **81**, 313, 445, 465 (1933).

[5] H. H. Nielsen, Phys. Rev. **38**, 1432 (1931).

$J=10$

1st equation

$+23531860b^3$ in the coefficient of W should be $+23457060b^3$.

2nd equation

$-23531860b^3$ in the coefficient of W should be $-23457060b^3$

-8903025 in the constant coefficient should be -893025.

4th equation

$+3177400320b^2$ in the coefficient of W should be $-3177400320b^2$

$-2697416640b^4$ in the coefficient of W should be $+2697416640b^4$.

We have also computed the equations which determine the levels of $J=11$. They are:

$$W^5-220W^4+W^3(16368-8976b^2)$$
$$-W^2(489280-1476288b^2)$$
$$+W(5395456-70684416b^2+10644480b^4)$$
$$-14745600+956421120b^2-73907520b^4=0.$$

$$W^6-220W^5+W^4(16368-17556b^2)$$
$$-W^3(489280-3329568b^2)$$
$$+W^2(5395456-203708736b^2+55226160b^4)$$
$$-W(14745600-4622346240b^2+674105184 0b^4)$$
$$-31629312000b^2+1834170624 00b^4-154378224 00b^6=0.$$

$$W^6-W^5(286\mp 66b)+W^4(28743\mp 18810b-11121b^2)$$
$$-W^3(1234948\mp 1878228b-2660724b^2\pm 463716b^3)$$
$$+W^2(21967231\mp 79628340b-206124534b^2$$
$$\pm 100549548b^3+20148975b^4)-W(128816766$$
$$\mp 1370208906b-5931201876b^2\pm 6483676716b^3$$
$$+3161391030b^4\mp 339022530b^5)+108056025$$
$$\mp 7131697650b-51015150945b^2\pm 12207753162 0b^3$$
$$+1078967146 95b^4\mp 3341458197 0b^5-217094377 5b^6=0.$$

The numerical values which were chosen for the moments of inertia are $h/8\pi^2cA=27.8055$, $h/8\pi^2cB=14.499724$, $h/8\pi^2cC=9.279276$. These figures were made to differ slightly from those given by Mecke since by so doing the parameter of asymmetry b which enters the algebraic equations becomes the simple number -0.164, thus facilitating the numerical work. It must be emphasized that the moments of inertia just given are not in any way more accurate than those found by Mecke. They simply represent a definite zeroth order approximation and will form a starting point for further calculations.

Our results appear in Table I in the following form. The fourth column gives the actual values for the energy levels as obtained from the observed lines while the first column gives the amount by which the observed level is lower than the computed level. Thus the energy levels as determined from the asymmetrical rotator formulae using the above moments of inertia are given by the sum of the numbers appearing in columns one and four of Table I.

An inspection of this table reveals that the differences recorded in column one are often large, ranging as high as 279.19 cm^{-1} for the level 11_{11} although they are much smaller for the lower levels of each J group. While these differences are in part due to the fact that the moments of inertia chosen are not the true moments, they represent principally the change in energy caused by the centrifugal forces of the rotation. That is, the water molecule is not completely rigid but is slightly distorted by its rotation. We intend to consider this point more exactly in a future paper but for the present we would like to calculate the order of magnitude of the stretching and its effect upon the levels.

The water molecule is, of course, an asymmetrical rotator and any exact theory of the centrifugal distortion must start from this point. However, the highest τ states of any J group correspond in fairly good approximation to a simple spinning of the molecule about the axis of least moment of inertia A as indicated in Fig. 2a. The rotational energy in this approximation is $h^2J^2/8\pi^2A$ according to which the energy of the level 10_{10} for example would be 2781 cm^{-1} whereas the exact formula gives the value 2902 cm^{-1}; an error of 4 percent. It is most convenient to specify the positions of the nuclei in the rotating molecule by means of the displacement coordinates q, x and y where q is the change in distance between the two hydrogen

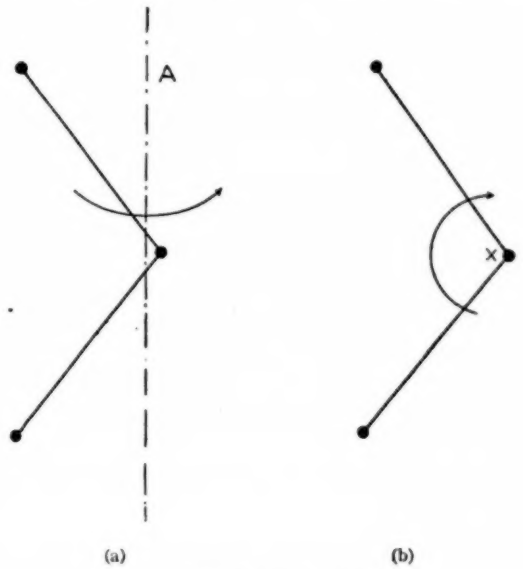


FIG. 2. The water molecule.

TABLE I. *The rotational energy levels of the water molecule.*

δW		J_τ	W	δW		J_τ	W	δW		J_τ	W	δW		J_τ	W
0	+ +	0_0	0	28.46	− +	6_5	1045.34	3.00	− −	8_{-6}	883.12	3.00	+ −	10_{-9}	1114.91
				12.83	− −	6_4	888.89	1.32	+ −	8_{-7}	744.31	3.00	+ +	10_{-10}	1114.91
0.01	+ −	1_1	42.30$_c$	12.83	+ −	6_3	888.85	1.32	+ +	8_{-8}	744.24				
.02	− −	1_0	37.06$_c$	5.05	+ +	6_2	757.84					279.19	+ −	11_{11}	3218.55
.00	− +	1_{-1}	23.78$_c$	4.92	− +	6_1	756.90	134.90	+ −	9_9	2226.41	279.19	− −	11_{10}	3218.55
				2.29	− −	6_0	661.63	134.90	− −	9_8	2226.41	193.68	− +	11_9	2974.88
.16	+ +	2_2	136.10	1.80	+ −	6_{-1}	649.11	84.29	− +	9_7	2010.67	193.68	+ +	11_8	2974.88
.19	− +	2_1	134.81	2.08	+ +	6_{-2}	602.78	84.29	+ +	9_6	2010.67	130.06	+ −	11_7	2741.42
.04	− −	2_0	95.04$_c$	1.30	− +	6_{-3}	553.00	49.64	+ −	9_5	1811.18	130.06	− −	11_6	2741.42
.04	+ −	2_{-1}	79.38$_c$	1.50	− −	6_{-4}	542.80	49.64	− −	9_4	1811.18	83.77	− +	11_5	2523.08
.05	+ +	2_{-2}	70.03$_c$	.51	+ −	6_{-5}	447.20	27.57	− +	9_3	1631.95	83.77	+ +	11_4	2523.08
				.54	+ +	6_{-6}	446.69	27.57	+ +	9_2	1631.81	52.51	+ −	11_3	2322.82
1.38	+ −	3_3	285.45					15.66	+ −	9_1	1477.78	52.47	− −	11_2	2322.77
1.37	− −	3_2	285.26	52.05	+ −	7_7	1395.26	15.47	− −	9_0	1475.46	33.80	− +	11_1	2145.17
.15	− +	3_1	212.07	52.05	− −	7_6	1395.26	11.65	− +	9_{-1}	1360.85	33.26	+ +	11_0	2144.06
.09	+ +	3_0	206.25	27.20	− +	7_5	1216.63	10.21	+ +	9_{-2}	1340.98	25.74	+ −	11_{-1}	2000.71$_c$
.01	+ −	3_{-1}	173.33	27.20	+ +	7_4	1216.63	11.40	+ −	9_{-3}	1283.29	26.65	− −	11_{-2}	1984.20
.05	− −	3_{-2}	142.17	12.80	+ −	7_3	1060.07	6.79	− −	9_{-4}	1216.62	27.39	− +	11_{-3}	1898.56
.06	− +	3_{-3}	136.74	12.80	− −	7_2	1059.88	8.48	− +	9_{-5}	1202.28	22.39	+ +	11_{-4}	1840.14
				5.75	− +	7_1	931.50	3.89	+ +	9_{-6}	1080.72	25.36	+ −	11_{-5}	1811.21
5.21	+ +	4_4	488.24	5.68	+ +	7_0	927.92	4.20	+ −	9_{-7}	1079.41	14.26	− −	11_{-6}	1694.71
5.21	− +	4_3	488.24	4.00	+ −	7_{-1}	842.65	2.06	− −	9_{-8}	920.41	14.64	− +	11_{-7}	1691.13$_c$
1.21	− −	4_2	383.90	3.06	− −	7_{-2}	816.78	2.06	− +	9_{-9}	920.41	7.68	+ +	11_{-8}	1525.65
1.29	+ −	4_1	382.49	3.76	− +	7_{-3}	782.54					7.80	+ −	11_{-9}	1525.36
.38	+ +	4_0	315.66	1.76	+ +	7_{-4}	709.54	198.44	+ +	10_{10}	2703.28	4.33	− −	11_{-10}	1327.58
.13	− +	4_{-1}	300.33	2.03	+ −	7_{-5}	704.40	198.44	− +	10_9	2703.28	4.33	− +	11_{-11}	1327.58
.52	− −	4_{-2}	275.21	.82	− −	7_{-6}	586.53	131.51	− −	10_8	2472.44				
.12	+ −	4_{-3}	224.74	.84	− +	7_{-7}	586.32	131.51	+ −	10_7	2472.44		− −	12_6	3033.78
.28	+ +	4_{-4}	221.90					83.25	+ +	10_6	2255.08		+ −	12_5	3033.78
				86.75	+ +	8_8	1789.77	83.25	− +	10_5	2255.08		− +	12_{-9}	1775.26$_c$
13.50	+ −	5_5	742.30	86.75	− +	8_7	1789.77	50.25	− −	10_4	2055.04		− −	12_{-10}	1775.26$_c$
13.50	− −	5_4	742.30	50.12	− −	8_6	1591.47	50.25	+ −	10_3	2055.04		+ −	12_{-11}	1558.42
4.92	− +	5_3	610.42	50.12	+ −	8_5	1591.47	29.79	+ +	10_2	1876.16		+ +	12_{-12}	1558.42
4.89	+ +	5_2	610.21	27.07	+ +	8_4	1411.89	29.75	− +	10_1	1875.68				
1.45	+ −	5_1	508.86	27.07	− +	8_3	1411.89	18.87	− −	10_0	1726.23		+ +	13_{-10}	2042.99
1.45	− −	5_0	503.94	13.57	− −	8_2	1256.24	18.56	+ −	10_{-1}	1719.75		+ −	13_{-11}	2042.99
.88	− +	5_{-1}	446.52	13.59	+ −	8_1	1255.45	20.53	+ +	10_{-2}	1614.43		− −	13_{-12}	1807.36$_c$
.55	+ +	5_{-2}	416.02	7.89	+ +	8_0	1132.10	13.75	− +	10_{-3}	1581.88		− +	13_{-13}	1807.36$_c$
.73	+ −	5_{-3}	399.45	7.13	− +	8_{-1}	1123.11	16.70	− −	10_{-4}	1538.65				
.32	− −	5_{-4}	326.49	4.65	− −	8_{-2}	1052.92	9.67	+ −	10_{-5}	1446.68		+ −	14_{-13}	2074.30
.40	− +	5_{-5}	325.23	4.52	+ −	8_{-3}	1006.38	11.22	+ +	10_{-6}	1438.50		+ +	14_{-14}	2074.30
				5.89	+ +	8_{-4}	983.22	5.57	− +	10_{-7}	1294.07				
28.46	+ +	6_6	1045.34	2.56	− +	8_{-5}	885.84	5.73	− −	10_{-8}	1293.49				

atoms. x and y are the displacements of the oxygen atom relative to the center of gravity of the hydrogens $\parallel$ and $\perp$ to the A axis, respectively. Since the centrifugal forces for this simple spinning motion are always perpendicular to the A axis, it is clear that x will be uninfluenced by them. We may accordingly omit it from the calculations. The total energy, rotational plus potential, has the form,

$$W/hc = hJ^2/8\pi^2cA + \tfrac{1}{2}(ay^2 + bq^2 + 2dqy),$$

where a, b and d are potential constants whose values will be considered later. A is the moment of inertia of the rotating molecule and is equal to $2mM(y_0+y)^2/(2m+M)$, where y_0 is the distance between the oxygen and the line joining the hydrogens in the nonrotating molecule. The nature of the motion is evidently such that the coordinates q and y will possess constant values which may be found by setting the derivatives of W to zero.

Thus $(1/hc)(\partial W/\partial q) = bq + dy = 0.$

Let this relation be used to eliminate q from the energy W and introduce the notation,

$\alpha = h/8\pi^2\mu y_0^2 c$, $\beta = (ab-d^2)y_0^2/2b$ and $\eta = y/y_0$,

then $W/hc = \alpha J^2/(1+\eta)^2 + \beta\eta^2.$

The value of η is now obtained by differentiation and is a root of the equation,

$$\eta(1+\eta)^3 = \alpha J^2/\beta.$$

The change in energy due to the rotational stretching is the difference between W/hc and αJ^2 and is readily found to be

$$\delta W/hc = -\beta\eta^2(1+3\eta+\eta^2).$$

The change in energy just calculated is a rigorous consequence of the assumptions employed, namely simple spinning about the A axis and a potential energy expression involving only quadratic terms. If $\delta W \ll W$, the formula may be developed and only first-order terms retained, in which case

$$(\delta W)_0/hc = -\beta\eta^2 = -\alpha^2 J^4/\beta.$$

The numerical value of α is of course 27.8055 cm^{-1} and the distance y_0 is known from Mecke's work to be 0.5845×10^{-8} cm. The potential energy constants a, b, d may be found from Bonner's[6] analysis of the vibrational bands of water. The coordinates used by Bonner are the valence coordinates but a transformation from his constants to ours may easily be effected, for example, through the formulae given by Sutherland and Dennison.[7] Combining these quantities we find $\beta = 37{,}635$ cm^{-1}. The numerical values of $(\delta W)_0$ and of δW may now easily be computed and are listed in the second and third columns of Table II in the units of cm^{-1}. The fourth column represents the difference between the calculated energy levels of the asymmetrical rotator and the observed levels of the water molecule. It is thus apparent that the principal cause of the difference between the computed and observed levels is due to the centrifugal force stretching.

It is interesting to note that the first-order calculation $(\delta W)_0$ differs considerably from the exact calculation δW particularly for the larger J values. This is important since it shows that no first-order computation of the stretching effect, as for example that proposed by Wilson,[8] can possibly be adequate for values of δW much above 5 cm^{-1}. A second way of showing this is to point out that for the higher energies of rotation the centrifugal forces are so great that the distortion of the molecule is no longer small. Thus for $J=11$, the bond angle changes from its equilibrium value of 104° 36′ to 98° 52′ while the O−H distance changes from 0.9558A to 0.9640A.

The differences between δW and δW_{obs} arise from two causes, the principal one of which is that δW has been calculated from much too simple a model. The system must eventually be treated wave mechanically and also the influence of anharmonic terms in the potential must be considered. The second cause of the difference is that the asymmetric rotator energy levels have been calculated using what might be called trial moments of inertia. If δW were known exactly, then $\delta W - \delta W_{obs}$ would immediately furnish the appropriate corrections to the trial moments. No significance is to be attributed to the fact that $(\delta W)_0$ appears to agree with the observed differences better than δW.

The values which have just been considered apply to the highest levels of each J group. Similar calculations may be made on the lowest levels of each group where the motion approximates to a simple spinning about the greatest moment of inertia axis as shown in Fig. 2b. Since this moment is larger, the angular velocity is lower and the corrections turn out to be sufficiently small so that only the first order term $(\delta W)_0$ need be considered. This together with δW_{obs} is shown in the fifth and sixth columns of Table II. The agreement seems to be rather less good than that for the high levels, possibly due to the fact that the moments of inertia of the molecule are such that the low levels correspond to simple spinning about a principal axis less well than the higher levels.

TABLE II. *Numerical values of* $(\delta W)_0$ *and of* δW.

J	$(\delta W)_0$	δW	δW_{obs}	$(\delta W)_0$	δW_{obs}
0	0	0	0	0	0
1	0.02	0.02	0.02	0.00	0.00
2	0.33	0.33	0.18	0.00	0.04
3	1.66	1.63	1.38	0.02	0.05
4	5.26	5.08	5.21	0.06	0.20
5	12.84	12.18	13.50	0.14	0.36
6	26.62	24.72	28.46	0.30	0.52
7	49.33	44.67	52.05	0.55	0.83
8	84.15	74.18	86.75	0.95	1.32
9	134.79	115.40	134.90	1.52	2.06
10	205.44	170.52	198.44	2.31	3.00
11	300.78	241.68	279.19	3.38	4.33

[6] L. G. Bonner, Phys. Rev. **46**, 458 (1934).
[7] G. B. B. M. Sutherland and D. M. Dennison, Proc. Roy. Soc. **148**, 250 (1935), see page 258.
[8] E. B. Wilson, Jr., J. Chem. Phys. **4**, 526 (1936).

TABLE III. *The rotation lines of water vapor.*

ν_{obs}	ν_{calc}	I	TRANSITION	ν_{obs}	ν_{calc}	I	TRANSITION	ν_{obs}	ν_{calc}	I	TRANSITION
554.97	555.00	(0.24)?	$7_4 - 6_0$	384.21				223.82	223.84	(50)	$8_{-3} - 7_{-3}$
550.33				380.45				221.74	221.88	3	$10_{-6} - 9_{-4}$
548.04				378.60	378.80	?	$8_{-1} - 8_{-7}$		221.78	(110)	$5_{-1} - 4_{-3}$
546.65	546.65	0.12	$11_{-4} - 10_{-8}$	376.57	376.50	12	$9_3 - 8_1$		217.36	.02	$10_8 - 10_6$
545.70	545.70	(0.06)	$11_{-2} - 10_{-6}$	375.54	375.57	3.9	$9_2 - 8_2$				$10_7 - 10_5$
536.48	536.48	(0.26)?	$6_5 - 5_1$	374.80	374.84	22	$8_6 - 7_4$		215.74	.03	$9_9 - 9_7$
526.23	526.27	0.79	$10_{-5} - 9_{-9}$				$8_5 - 7_5$				$9_8 - 9_6$
519.81				370.26	370.09	2.5	$7_{-2} - 6_{-6}$	215.31	215.23	(1.4)?	$3_2 - 2_{-2}$
518.09	518.09	0.26	$10_{-6} - 9_{-8}$		369.77	0.6	$11_{-2} - 10_{-2}$	214.70	214.66	18	$10_{-7} - 9_{-7}$
	517.47	(0.1)?	$10_{-1} - 9_{-5}$	358.90	358.90	(6.2)	$10_{-1} - 9_{-1}$	214.00	214.00	23	$7_{-2} - 6_{-2}$
[517.14	517.14	0.34	$11_{-5} - 10_{-7}$	357.45	357.45	(10)	$6_1 - 5_{-3}$	212.68	212.77	6	$10_{-8} - 9_{-6}$
515.27	515.27	0.04	$11_{11} - 10_9$	354.67	354.67	18	$9_1 - 8_{-1}$		212.67	25	$11_{-11} - 10_{-9}$
			$11_{10} - 10_{10}$	354.47	354.47	(15)	$9_{-1} - 8_{-3}$				$11_{-10} - 10_{-10}$
[510.70	510.70	0.07	$12_5 - 11_5$	351.88	352.01	9	$8_4 - 7_2$	208.53	208.53	260	$5_1 - 4_{-1}$
			$12_6 - 11_4$		351.82	27	$8_3 - 7_3$	203.12	202.98	110	$4_4 - 3_2$
507.10	507.07	?	$7_3 - 6_{-3}$	349.92	349.92	44	$7_7 - 6_5$	202.79	202.79	330	$4_3 - 3_3$
502.45	502.47	(0.7)?	$10_{-3} - 9_{-7}$				$7_6 - 6_6$	202.55	202.59	(140)	$6_{-1} - 5_{-1}$
	502.44	0.10	$11_9 - 10_7$	346.06	346.07	?	$4_4 - 3_{-2}$	200.44	200.04	0.05	$10_6 - 10_4$
			$11_8 - 10_8$	343.38	343.38	(16)	$8_{-2} - 7_{-4}$				$10_5 - 10_3$
492.22	492.24	(0.4)	$9_0 - 8_{-4}$		343.36	5.5	$9_0 - 8_0$	199.31	199.49	0.1	$9_7 - 9_5$
486.34	486.34	0.24	$11_7 - 10_5$	335.34	335.34	6.5	$7_{-3} - 6_{-5}$				$9_6 - 9_4$
			$11_6 - 10_6$		335.00	(0.4)?	$5_2 - 4_{-2}$				$8_8 - 8_6$
484.18				328.34	328.32	13	$8_2 - 7_0$		198.30	0.2	$8_7 - 8_5$
481.28							$7_5 - 6_3$		198.07	0.3	$11_{-8} - 11_{-10}$
478.05				327.78	327.77	74	$7_4 - 6_4$	197.62	197.74	0.9	$11_{-9} - 11_{-11}$
476.87	476.87	0.30	$10_{10} - 9_8$	323.90	323.95	39	$8_1 - 7_1$		197.67	(0.8)	$10_{-4} - 9_{-2}$
			$10_9 - 9_9$		323.88	(6)	$6_{-1} - 5_{-5}$		197.60	12	$9_{-6} - 8_{-6}$
472.68	472.38	0.62	$9_{-4} - 8_{-8}$	315.32	315.32	19	$8_0 - 7_{-2}$	195.88	195.90	(9)	$9_{-5} - 8_{-3}$
468.88					314.91	(1.6)?	$4_3 - 3_{-1}$	194.47	194.50	50	$10_{-9} - 9_{-9}$
468.04	468.04	0.56	$11_5 - 10_3$	305.00							$10_{-10} - 9_{-8}$
			$11_4 - 10_4$	303.30	303.17	80	$7_3 - 6_1$	193.57	193.57	35	$9_{-7} - 8_{-5}$
461.77	461.77	0.75	$10_8 - 9_6$	303.00	303.04	135	$6_6 - 5_4$	188.28	188.28	68	$5_0 - 4_0$
			$10_7 - 9_7$				$6_5 - 5_5$	183.58	183.63	?	$5_1 - 5_{-5}$
458.05	457.97	1.78	$9_{-5} - 8_{-7}$	302.04	302.04	27	$7_2 - 6_2$	181.46	181.44	67	$8_{-5} - 7_{-5}$
	457.93	(1.0)?	$10_{-4} - 9_{-6}$		301.49	0.7	$11_{-4} - 10_{-4}$		179.58	0.7	$8_6 - 8_4$
	457.86	(0.45)	$9_{-2} - 8_{-6}$	298.61	298.59	(6)	$10_{-3} - 9_{-3}$				$8_5 - 8_3$
	457.10	(0.1)?	$7_2 - 6_{-2}$	290.91					179.37	0.1	$9_4 - 9_2$
453.96				289.65	289.65	140	$7_{-1} - 6_{-3}$		179.36	0.1	$10_3 - 10_1$
451.88	451.88	0.30	$11_{-3} - 10_{-5}$		288.06	(2)	$9_{-2} - 8_{-2}$		179.23	0.3	$9_5 - 9_3$
447.14	447.14	0.84	$11_3 - 10_1$	282.45	282.39	102	$7_1 - 6_{-1}$	179.12	179.16	1.3	$10_{-7} - 10_{-9}$
446.61	446.61	0.28	$11_2 - 10_2$		282.04	(3)	$5_0 - 4_{-4}$		178.88	0.03	$10_4 - 10_2$
443.90	443.90	1.8	$10_6 - 9_4$	280.46	280.46	42	$8_{-1} - 7_{-1}$	178.62	178.63	1	$7_7 - 7_5$
			$10_5 - 9_5$	278.55	278.68	47	$6_4 - 5_2$				$7_6 - 7_4$
442.24	442.33	(4.0)?	$6_3 - 5_{-1}$		278.43	142	$6_3 - 5_3$		178.58	0.4	$10_{-8} - 10_{-10}$
436.64	436.64	2.0	$9_9 - 8_7$	276.29	276.29	(35)	$6_{-2} - 5_{-4}$	177.65	177.65	100	$4_2 - 3_0$
			$9_8 - 8_8$	267.73	267.73	1.5	$13_{-11} - 12_{-9}$	176.12	176.12	90	$9_{-8} - 8_{-8}$
434.98							$13_{-10} - 12_{-10}$				$9_{-9} - 8_{-7}$
431.35	431.35	?	$6_2 - 5_{-4}$	266.94	266.94	1.5	$14_{-14} - 13_{-12}$	173.50	173.58	20	$8_{-6} - 7_{-4}$
429.09	429.00	?	$10_1 - 10_{-5}$				$14_{-13} - 13_{-13}$		173.49	(50)	$4_0 - 3_{-2}$
426.42	426.64	(0.1)?	$5_4 - 4_0$	266.29	266.29	30	$7_0 - 6_0$	170.40	170.42	270	$4_1 - 3_1$
425.42	425.42	1.5	$11_1 - 10_{-1}$	256.21	256.21	1.4	$11_{-6} - 10_{-6}$	169.56	169.06	0.1	$11_{-6} - 11_{-8}$
423.14	423.14	3.4	$10_4 - 9_2$	254.06	254.06	300	$5_5 - 4_3$	166.78	166.74	35	$7_{-4} - 6_{-4}$
			$10_3 - 9_3$				$5_4 - 4_4$	166.32	166.44	(7)	$8_{-4} - 7_{-2}$
420.10	420.06	3.9	$8_{-3} - 7_{-7}$		253.90	57	$6_2 - 5_0$	161.90	162.00	?	$4_2 - 4_{-4}$
419.20	419.20	4.6	$9_7 - 8_5$		249.90	3.2	$12_{-9} - 11_{-9}$	160.21	160.31	0.9	$9_{-6} - 9_{-8}$
			$9_6 - 8_6$		249.61	1.1	$12_{-10} - 11_{-8}$		160.18	(3)	$9_{-3} - 8_{-1}$
418.68	418.71	(2.2)	$8_{-1} - 7_{-5}$		248.94	4.3	$13_{-13} - 12_{-11}$	158.96	159.00	2.7	$9_{-7} - 9_{-9}$
417.88	417.83	0.5	$11_0 - 10_0$				$13_{-12} - 12_{-12}$	158.00	157.99	107	$8_{-7} - 7_{-7}$
400.70	400.70	1.5	$10_2 - 9_0$	248.01	248.04	160	$6_1 - 5_1$	157.70	157.71	36	$8_{-8} - 7_{-6}$
399.29	399.29	9.1	$9_5 - 8_3$	245.65	245.75	(9)	$4_1 - 3_{-3}$		156.75	0.8	$7_4 - 7_2$
			$9_4 - 8_4$		245.61	(60)	$6_0 - 5_{-2}$		156.56	2.4	$7_5 - 7_3$
397.81	397.90	4.4	$10_1 - 9_1$	244.46	244.40	11	$10_{-5} - 9_{-5}$	156.49	156.47	4	$6_6 - 6_4$
	397.81	(2.2)	$10_{-2} - 9_{-4}$	233.40	233.40	(6.5)	$9_{-4} - 8_{-4}$				$6_5 - 6_3$
	397.45	(2.5)?	$9_{-3} - 8_{-5}$	232.16	232.16	3	$11_{-8} - 10_{-8}$		156.44	1.3	$8_3 - 8_1$
396.79	396.69	1.2	$8_{-4} - 7_{-6}$	231.29	231.29	8	$11_{-9} - 10_{-7}$		156.35	0.2	$9_2 - 9_0$
394.51	394.51	11	$8_8 - 7_6$	230.84	230.84	11	$12_{-12} - 11_{-10}$	155.62	155.93	0.2	$10_1 - 10_{-1}$
			$8_7 - 7_7$				$12_{-11} - 11_{-11}$		155.65	0.4	$8_4 - 8_2$
385.58	385.68	?	$5_2 - 4_{-3}$		229.33	(0.8)	$11_{-5} - 10_{-3}$		154.17	0.5	$9_3 - 9_1$
385.17	385.25	(1.6)	$10_0 - 9_{-2}$	227.93	227.93	250	$5_3 - 4_1$	153.55	153.55	(130)	$6_{-3} - 5_{-3}$
	385.12	(1.8)	$7_0 - 6_{-4}$	226.38	226.31	83	$5_2 - 4_2$	152.55	152.61	0.8	$10_{-5} - 10_{-7}$

TABLE III.—*Continued.*

ν_{obs}	ν_{calc}	I	TRANSITION	ν_{obs}	ν_{calc}	I	TRANSITION	ν_{obs}	ν_{calc}	I	TRANSITION
151.37	151.40	88	$7_{-5} - 6_{-3}$		124.36	0.4	$9_{-2} - 9_{-4}$		99.04	40	$2_2 - 1_0$
150.62	150.64	290	$3_3 - 2_1$		124.12	0.7	$8_2 - 8_0$		98.84	(6)	$6_{-2} - 5_0$
	149.93	0.1	$10_2 - 10_0$	122.95	123.01	3	$7_{-4} - 7_{-6}$		96.21	3	$6_2 - 6_0$
149.15	149.16	95	$3_2 - 2_2$		122.87	1.5	$9_{-5} - 9_{-7}$	96.12	96.11	(17)	$6_{-1} - 6_{-3}$
144.81	145.01	0.3	$10_{-6} - 10_{-8}$	121.97	121.97	180	$6_{-5} - 5_{-5}$		96.11	4	$6_{-4} - 6_{-6}$
141.50	141.53	5.4	$8_{-5} - 8_{-7}$		120.54	3.7	$8_{-2} - 8_{-5}$		94.42	(0.03)	$10_0 - 9_2$
140.81	140.81	65	$5_{-2} - 4_{-2}$	120.20	120.20	60	$6_{-6} - 5_{-4}$	92.51	92.51	124	$2_1 - 1_1$
139.84	139.84	50	$7_{-6} - 6_{-6}$	118.08	118.08	8	$7_{-3} - 7_{-7}$	89.53	89.53	6	$5_{-2} - 5_{-4}$
139.09	139.12	150	$7_{-7} - 6_{-5}$		116.93	0.7	$9_1 - 9_{-1}$	88.42	88.85	5	$7_1 - 7_{-1}$
	138.97	(0.1)	$10_{-2} - 9_0$	116.74	116.73	4	$8_{-1} - 8_{-3}$	88.00	88.00	165	$4_{-3} - 3_{-3}$
	138.88	1.7	$8_{-6} - 8_{-8}$		111.80	0.3	$10_0 - 10_{-2}$		87.92	9	$5_0 - 5_{-2}$
138.09	137.87	0.4	$10_{-1} - 10_{-3}$	111.21	111.21	51	$3_0 - 2_0$		85.83	(0.04)	$10_{-5} - 9_{-1}$
	135.90	0.6	$9_{-4} - 9_{-6}$		111.18	2	$7_0 - 7_{-2}$	85.62	85.75	(3)	$7_{-1} - 6_1$
	135.20	0.4	$10_{-3} - 10_{-5}$	107.92	107.79	13	$6_1 - 6_{-1}$		84.75	0.02	$9_{-2} - 8_2$
	134.48	0.3	$9_0 - 9_{-2}$	107.30	107.24	2.4	$7_{-2} - 7_{-4}$		84.51	(0.04)	$9_{-4} - 8_0$
133.45	133.43	(20)	$7_{-3} - 6_{-1}$		106.27	6	$5_2 - 5_0$	82.16	82.16	30	$4_1 - 4_{-1}$
132.69	132.69	(150)	$3_1 - 2_{-1}$	105.78	105.80	13	$6_{-3} - 6_{-5}$		81.01	0.6	$9_{-3} - 9_{-5}$
	132.34	2.4	$8_1 - 8_{-1}$		105.75	17	$4_3 - 4_1$	79.59	79.73	45	$4_{-4} - 3_{-2}$
	132.09	3.0	$5_4 - 5_2$		105.40	(0.5)	$9_{-1} - 8_1$		79.18	0.5	$8_0 - 8_{-2}$
131.89	131.96	1.6	$7_2 - 7_0$	104.59	104.59	60	$5_{-4} - 4_{-4}$	79.02	79.01	8	$3_2 - 3_0$
	131.95	7.9	$6_3 - 6_1$		104.34	5	$4_4 - 4_2$	78.26	78.14	4	$7_{-3} - 7_{-5}$
	131.88	8.8	$5_5 - 5_3$		104.10	0.07	$10_{-3} - 9_1$		77.56	0.5	$9_{-1} - 9_{-3}$
130.90	131.05	2.6	$6_4 - 6_2$	101.46	101.56	16	$5_3 - 5_1$		75.78	0.04	$10_{-2} - 10_{-4}$
128.64	128.57	4.6	$7_3 - 7_1$	100.59	100.49	180	$5_{-5} - 4_{-3}$	75.59	75.59	20	$4_{-1} - 4_{-3}$
126.96	127.00	(160)	$4_{-1} - 3_{-1}$		100.15	0.1	$10_{-4} - 10_{-6}$		74.88	0.7	$8_{-3} - 7_1$
126.61	126.78	35	$6_{-4} - 5_{-2}$		100.10	1	$8_{-4} - 8_{-6}$	74.22	74.22	13	$5_{-3} - 5_{-5}$
	125.00	(1)	$8_{-2} - 7_0$	99.04	99.12	96	$5_{-3} - 4_{-1}$				

2. The intensities of the rotation lines

The first step towards making an analysis is now complete. We are furnished with the energy levels of an asymmetrical molecule whose moments of inertia closely approximate those of water. The order of magnitude of the centrifugal stretching has been calculated for the highest and lowest levels and presumably will lie between these two limits for the intervening levels. The selection rules governing the allowed transitions of the asymmetric rotator are well known.[9] Each level is given a symbol, $(++)$, $(+-)$, $(-+)$ or $(--)$, which denotes the symmetry group to which the level belongs. If the electric moment lies along the middle axis of inertia, as it does in water vapor, the only radiative transitions are those connecting the levels $(++) \longleftrightarrow (--)$ or $(+-) \longleftrightarrow (-+)$. In addition to this $\Delta J = \pm 1$ or 0. The formula for the intensity of the line corresponding to the transition from the state J_τ'' to the state $J_{\tau'}$ may be readily obtained.

$$I_{J',J''_{\tau',\tau''}} = (8\pi^3 \nu N_{J''_{\tau''}}/3hc) \times (1 - e^{-h\nu/kT})(M_{J',J''_{\tau',\tau''}})^2.$$

In this expression ν is the frequency of the line and $N_{J''_{\tau''}}$ is the number of molecules per cc in the initial state. It has the form

$$N_{J\tau} = (N g_{J\tau} e^{-W_{J\tau}/kT}) / \sum g_{J\tau} e^{-W_{J\tau}/kT},$$

where N is the total number of water molecules and $g_{J\tau}$, the weight of the state, is equal to $(2J+1)f$. The factor f is due to the existence of the hydrogen nuclear spin of $\frac{1}{2}(h/2\pi)$ and is equal to 3 for τ odd and 1 for τ even.

The quantity $M_{J',J''_{\tau',\tau''}}$ is the matrix element of the electric moment associated with the transition and may be calculated from the minors of the secular energy determinant of the asymmetric rotator. Such calculations were made earlier by one of us (D.M.D.) for the matrix elements connecting all the states up to and including $J=3$. The resulting expressions were algebraic functions of the asymmetry parameter b whose complexity increased so rapidly with J that it seemed inadvisable to attempt to extend them to higher J values. Moreover a study of those matrix elements which were determined revealed that they could usually be satisfactorily approximated in the following manner. The levels of an asymmetric rotator may be divided

[9] See, for example, D. M. Dennison, Rev. Mod. Phys. **3**, 280 (1931).

roughly into two main classes and an intervening region. The first class in any J group are characterized by large τ values and the levels tend to associate themselves in pairs. The association becomes less marked as τ decreases. This class of levels corresponds to a rotation of the molecule about the axis of least moment of inertia together with the accompanying precession and nutation. The energies and wave functions are quite well approximated by symmetrical rotator energies and wave functions. Thus for water vapor the levels of $J=10$ from $\tau=10$ to $\tau=-1$ belong to the first class. The second class of levels extends upwards from the lowest τ values and corresponds to rotations about the greatest moment of inertia axis. In the present example they would include the levels from $\tau=-10$ to $\tau=-5$. The region between these classes corresponds roughly to rotations about the middle axis and is not well approximated by symmetrical rotator wave functions. The point of this discussion is that a matrix element connecting two levels belonging to the same class was found to differ very little from the corresponding matrix element of a symmetrical rotator. The numbering of the asymmetrical and symmetrical rotator levels is of course quite different and care must be used in selecting the proper indices. Thus in the first class of levels the indices $\tau=J$ and $J-1$ are to be identified with $K=J$ of the symmetric rotator and $\tau=J-2$ and $J-3$ with $K=J-1$ etc. In the second class $\tau=-J$ and $-J+1$ becomes $K=J$. When a transition occurs from a level of one class to that of another or to the intervening region, we have calculated the intensity as from each class separately and taken an average value. The approximation appears to be somewhat rough but it has proved a guide in identifying many lines and it will be seen later that our calculated intensities are usually in good agreement with those observed.

The formulae governing the intensities may now be given in explicit form providing we adopt the approximation of using the matrix elements derived from symmetric rotator wave functions.[9] It is convenient to divide them into three classes.

$$I_J{}^{J-1}{}_K{}^{K-1}=(CF\nu/c)e^{-W_{J-1,\,K-1}/kT}(1-e^{-h\nu/kT})\times(J+K)(J+K-1)/J, \quad \text{(a)}$$

$$I_J{}^{J-1}{}_K{}^{K+1}=(CF\nu/c)e^{-W_{J-1,\,K+1}/kT}(1-e^{-h\nu/kT})\times(J-K)(J-K-1)/J, \quad \text{(b)}$$

$$I_J{}^{J}{}_K{}^{K-1}=(CF\nu/c)e^{-W_{J,\,K-1}/kT}(1-e^{-h\nu/kT})\times(2J+1)(J+K)(J-K+1)/J(J+1). \quad \text{(c)}$$

The factor F is equal to $\frac{1}{4}$ or $\frac{3}{4}$ according as the subscripts τ are even or odd, respectively. The constant

$$C=4\pi^3NM_0{}^2/3h\sum g_{J\tau}e^{-W_{J\tau}/kT},$$

where M_0 is the permanent electric moment of the water molecule. In the present paper no attempt will be made to evaluate the absolute intensities of the lines although a preliminary calculation shows them to be of the right order of magnitude. The numbers appearing in Table III under the heading I are in reality the absolute intensities divided by the constant C. It may be of value to give one or two examples of these calculations in detail. Consider the line $5_3-5_1=101.46$. A study of the levels shows that both the initial and final states belong to one class of levels, namely, those corresponding to rotations about the least moment of inertia axis. Since $\Delta J=0$ the formula (c) must be used with $J=5$, $K=4$ and $F=\frac{3}{4}$. The resulting substitution yields $I=16.2C$.

The majority of the lines correspond to transitions from a level of one class to a level of the same class and for these our approximation should be well justified. These intensities are listed without braces. In the remaining cases, where the numbers are enclosed with braces, an interpolation of uncertain validity has been made. The line $6_{-2}-5_{-4}=276.29$ is such a one. If this is considered to be a transition between two levels belonging to rotations about the least moment of inertia axis then formula (a) should be employed with $J=6$, $K=2$ and $F=\frac{1}{4}$ giving $I=96.4C$. Such an interpretation would require the levels 6_{-2}, 6_{-3} and 5_{-3}, 5_{-4} to form close pairs. This is obviously false, particularly in the latter case. If we proceed under the assumption that the levels belong to the class of rotations about the greatest axis then formula (b) is applicable with $J=6$, $K=4$, $F=\frac{1}{4}$ and $I=3.4C$. This requires that 6_{-1}, 6_{-2} and 5_{-4}, 5_{-5} form pairs, a condition which is well met in the second

case but badly fulfilled in the first. Thus neither calculation is satisfactory although the latter is apparently to be preferred. Accordingly this intensity is listed as $\frac{1}{3}[96.4+2(3.4)]C=(34)C$. It is clear that the intensities in braces are merely to be regarded as estimates which in extreme cases may be in error by factors as high as ten. In a few cases the symmetric rotator approximation is completely uncertain and these are denoted by placing a question mark behind the brace.

3. The rotational energy levels

It is now possible to construct the energy levels of the water molecule beginning with $J=0$ and working successively towards higher J values. The results are listed in Table I. The six levels, 1_1, 1_0, 1_{-1}, 2_0, 2_{-1} and 2_{-2}, are followed by the letter c which means that they have been calculated using the moments of inertia given by Mecke.[4] A correction for the centrifugal stretching has been applied which was based on Wilson's[8] equations. The necessity for calculating these levels lies in the fact that the lines which would fix their values fall in the far infrared outside the region experimentally measured. It seems impossible, however, that our estimates could be in error by more than a very few hundredths of waves per-cm.

All the subsequent levels are derived from the observed lines by means of the following procedure. The position of an unknown level is approximately determined by subtracting a suitable correction from its computed value. The lines which connect this level with levels already known are then calculated and compared with the observed lines in the spectrum. For example, consider that point where the levels up through $J=5$ have been obtained. The two levels 6_6 and 6_5, as calculated from the asymmetric rotator formula, are practically coincident at 1073.76 cm^{-1}. The δW for $J=6$ was 24.72 and hence our first estimate for their position is 1049. There are only two transitions to $J=5$ levels, namely, a strong line $\left\{\begin{matrix}6_6-5_4\\6_5-5_5\end{matrix}\right\}$ and a weak line 6_5-5_1. Our initial estimates of these lines are accordingly 307 and 540 cm^{-1}. Turning to the spectrum we find a very strong line at 303.00 cm^{-1} and a weak line at 536.48, cm^{-1}. Identifying these with the transitions in question we obtain $\left\{\begin{matrix}6_6\\6_5\end{matrix}\right\}=1045.30$ and 1045.34, respectively. In this particular case it was felt that the weak line could be measured with slightly greater experimental accuracy and therefore the latter value was given preference. The next levels 6_4 and 6_3 are calculated to lie at 901.72 and 901.68. The δW is here estimated by observing that the δW of a next highest pair of one J group is about equal to the δW of the highest pair of the group $J-1$. This would put 6_4 and 6_3 at about 888 cm^{-1}. There are now three lines which correspond to the transitions, $\left\{\begin{matrix}6_6-6_4\\6_5-6_3\end{matrix}\right\}$, $\left\{\begin{matrix}6_4-5_2\\6_3-5_3\end{matrix}\right\}$ and 6_3-5_{-1}. These lines may easily be identified with those observed at, 156.49, 278.55 and 442.24 cm^{-1}.

In theory each level may be determined by means of two or three lines but in many cases (nearly $\frac{2}{3}$ of the levels) some of the lines are either too weak, are overlayed by much stronger lines, or fall outside the region investigated. This is particularly true for the high J groups and in fact none of the lines to the levels 11_{-1} and 11_{-7} were satisfactorily observable and consequently these values are estimated only. The fact that a level can be fixed by each of several lines is equivalent to the existence of combination relations. We might give, as examples, the following:

$8_{-7}-7_{-7}=158.00$	$7_{-5}-7_{-7}=118.08$
$8_{-5}-8_{-7}=141.50$	$8_{-5}-7_{-5}=181.46$
299.50	299.54
and $7_{-3}-6_{-5}=335.34$	$7_{-7}-6_{-5}=139.09$
$8_{-3}-7_{-3}=223.82$	$8_{-3}-7_{-7}=420.10$
559.16	559.19

The errors of 0.04 and 0.03, respectively, are to be distributed among four lines. This order of magnitude for the error is in accordance with what is revealed by a study of the experimental curves themselves, namely, that a sharp single line may be located to within 0.05 cm^{-1}.

The methods which have been outlined proved very satisfactory for identifying the lines and determining the energy levels. However, an additional process for correlating the lines was found which served as an important check on the

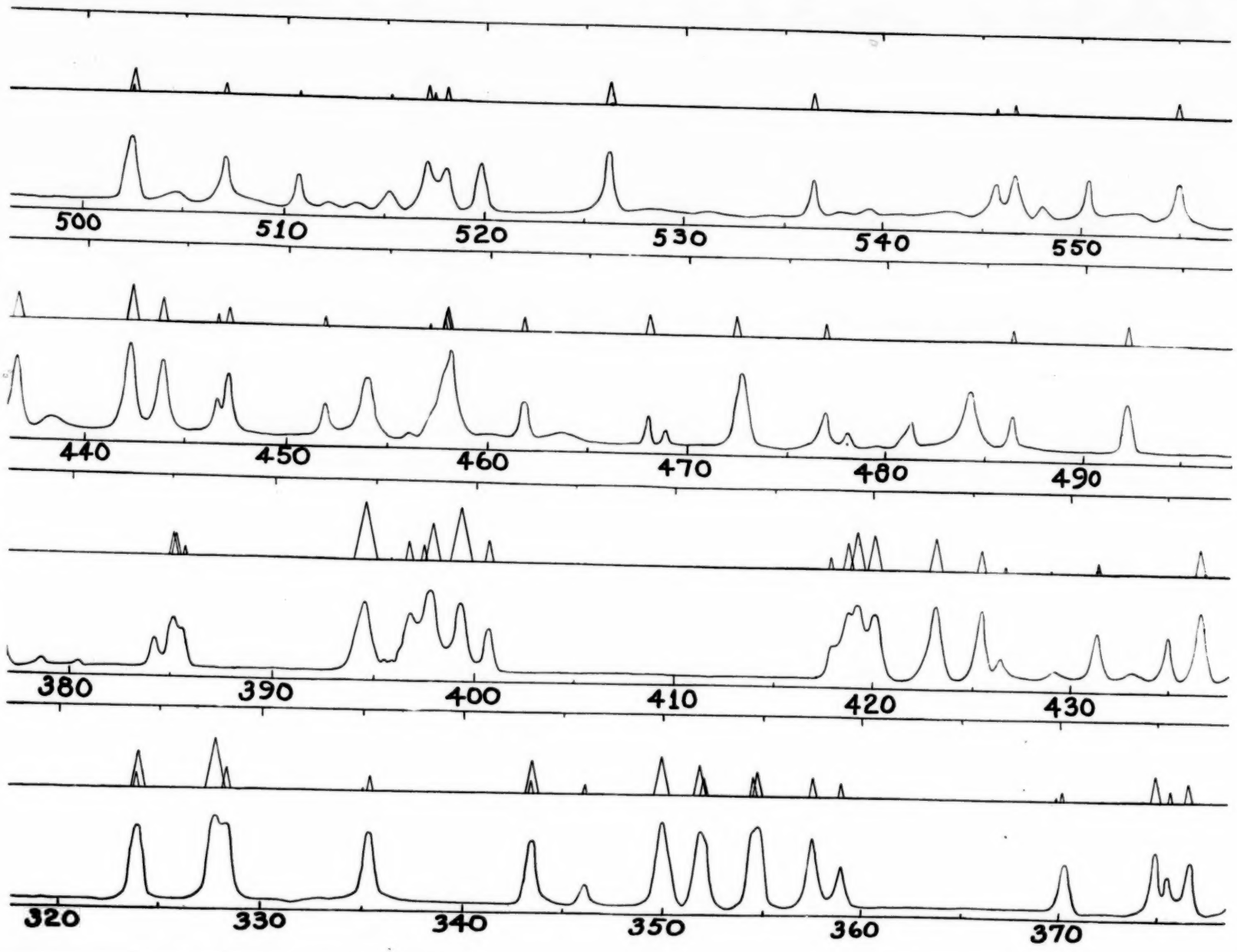


FIG. 3. Analysis of the rotational spectrum of water vapor.

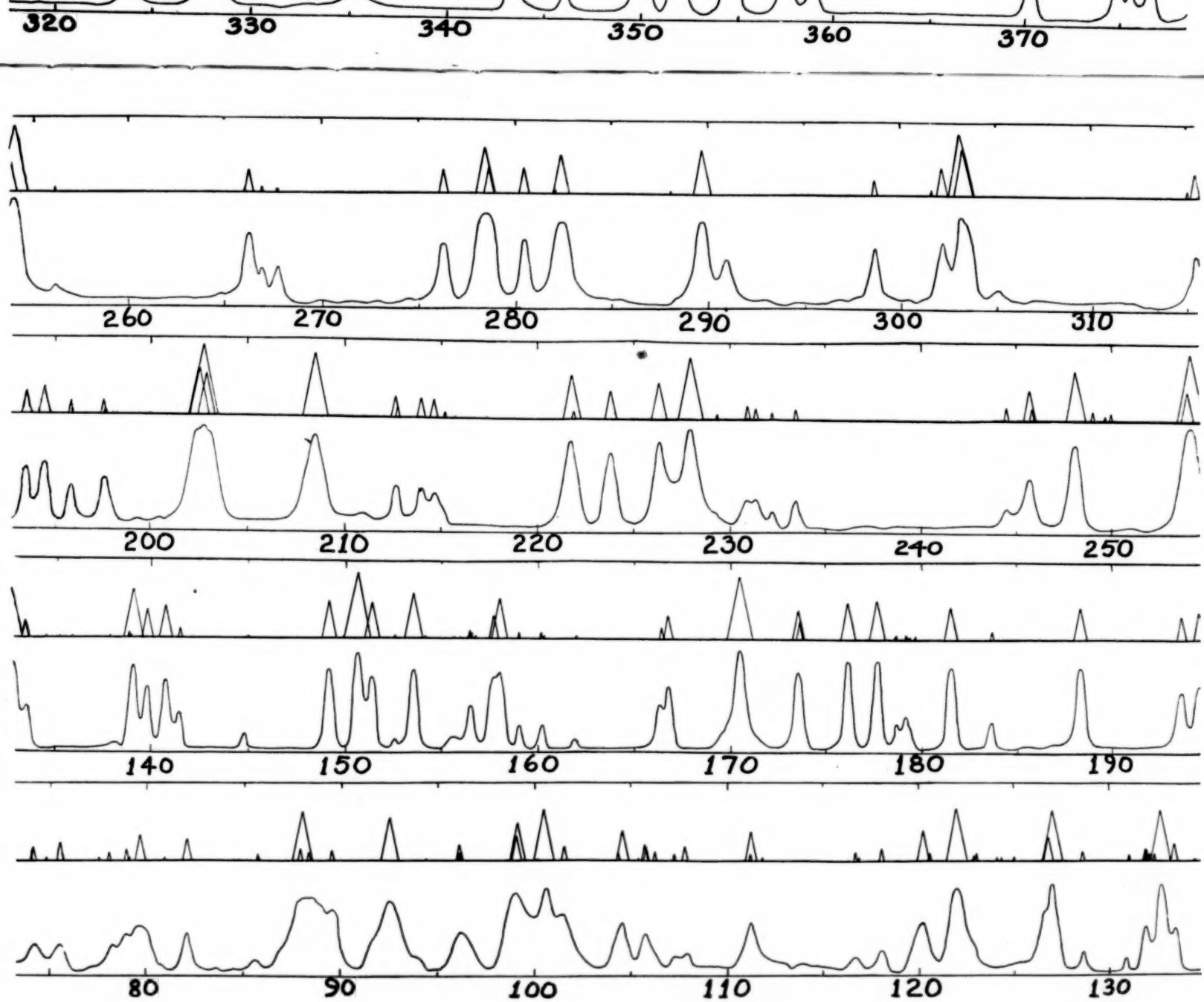


FIG. 4. Analysis of the rotational spectrum of water vapor.

identifications. This consisted in recognizing that many of the lines may be formed into series which are analytic in nature. As one example consider the transitions which connect the highest levels in one J group with the highest levels in the next lowest J group. The first observed members of this series are the two lines $2_1-1_1=92.51$ and $2_2-1_0=99.04$. An average value of these numbers, namely 95.78, is taken. There are ten members of the series of which the highest is $\left\{\begin{matrix}11_{11}-10_9\\11_{10}-10_{10}\end{matrix}\right\}=515.27$. The ten numbers are placed in a column and first and second differences are then taken. The second differences are found to have the values 1.03, 1.96, 2.17, 2.02, 2.33, 2.46, 1.90 and 1.83. The run of these numbers leaves no doubt but that the ten lines belong together and form a series. Second differences are of course very sensitive to changes in the frequencies of the original lines. For example the value 2.02 seems to be somewhat too low. It depends among others upon the line at 303.00 which is seen to be one component of a close doublet and hence rather inaccurately determined. If we should arbitrarily increase it to 303.08 then the third, fourth and fifth second differences would read 2.09, 2.18, 2.25; the others remaining unchanged. From the complete spectrum, in all six independent series were located, the longest containing ten numbers and the shortest, six. Other regularities exist between the energy levels which are revealed by a little inspection. The δW of one J group seems to be closely correlated with the δW of the next higher J group. Again from the highest level of a group to the lowest level δW decreases in a rather smooth manner except for a certain alternation. Thus starting from the bottom the δW of the 1st, 3rd etc. levels are high compared with the δW of the 2nd, 4th, etc., levels.

Through the use of the foregoing methods the rotational energy levels of water have been determined up through $J=11$. As a final test of our analysis we shall calculate the lines corresponding to every allowed transition together with the approximate expressions for the intensities. These are listed in Table III. A few of the lines connecting high energy states are not included in the table since their calculated intensities are so small as to show that they lie well below the limit of observation. There are certain transitions which are possible for the asymmetric rotator but which are forbidden in the corresponding symmetric rotator. The approximate intensity formulae are consequently not applicable to these lines and we have treated them as follows. In every case the frequency of the line was calculated and a search was made to see if it had been found experimentally. In the great majority of instances no line was observed showing that its intensity must be very small. Such transitions of which an example is $5_4-6_{-6}=295.61$ were omitted from the table. In a few cases, as $4_4-3_{-2}=346.07$, lines were observed and these are listed in the table with question marks for their intensities. These recorded lines are usually between levels of low J values where the asymmetric nature of the rotation is most pronounced and which are hence the most likely to appear.

The data contained in Table III are shown graphically in Figs. 3 and 4. The lower curve represents the observed spectrum of water vapor as obtained by averaging the points found from several independent runs over each line. Unfortunately the intensities of the lines are not determined with anything like the accuracy attained for their positions. The relative intensities of neighboring lines or groups of lines can usually be relied upon but the curve gives no measure of the strength of a line in one frequency range with that in another range. This is evident from the manner in which the data were obtained. For the region from 74 cm^{-1} to about 300 cm^{-1} the spectrometer case was kept as dry as possible with P_2O_5. From this point onward the dryer was gradually removed until for the highest frequencies open trays of water were placed in the spectrometer. Thus there can be no direct comparison between two lines at say 200 and 400 cm^{-1}.

Above the experimental curve we have plotted the calculated positions and intensities of the lines. These are drawn so that the area under each triangle is proportional to the intensities appearing in the table. Since the intensities in the experimental curve were being augmented in going toward shorter wave-lengths by admitting more water into the path, it was necessary to increase the theoretical intensities. This was done

somewhat arbitrarily by letting the proportionality factor be 1 from 74 cm^{-1} to 295 cm^{-1}, 2 from 295 to 380, 20 from 380 to 465 and 80 from 465 to 560 cm^{-1}.

4. Discussion

The figures which we have obtained for the rotational energy, in general, agree well with the values published by Mecke,[4] particularly for the lower levels of each J group. In the case of the higher levels there exist differences occasionally as large as 1 or 2 cm^{-1}. It appears to us that the present measurements of the far infrared spectrum allow a very much more precise determination of the levels than is attainable from any study of the existing data on the near infrared bands.

A comparison of our results with those found by Barnes, Benedict and Lewis[3] is less satisfactory since it shows a lack of agreement of from 1 to 9 cm^{-1}. A second and a principal difference between the present investigation and those which have gone before it, is the completeness with which the energy levels have been obtained. Thus it will be possible now to determine such quantities as for example the Zustandsumme at room temperature with an accuracy of 0.1 percent or better.

Let us now consider the degree of completeness with which the rotational spectrum of water vapor has been analyzed and the accuracy with which the levels have been determined. Two points of view must be presented. The first of these might be termed, the verification of the level scheme through combination relations. Of the total of 173 lines, 160 are satisfactorily accounted for. From these 160 lines, 115 distinct levels have been calculated thus implying the existence of 45 combination relations. Thus in principle there should be 45 predicted lines which may be compared with measured lines. Actually these lines are not isolated since each level was found by averaging the values as determined from all the lines ending upon the level in question. Nevertheless the exactness with which the combination relations are fulfilled is evinced by the good agreement between the observed and calculated frequencies in Table III. The smallness of the differences is an index of the uncertainty to be ascribed to the energy levels. For the levels up through $J=7$ or 8, where most of the combination relations occur, the order of magnitude of the error almost certainly is less than 0.1 cm^{-1} while for no level will it be more than 0.4 cm^{-1}.

The second method of testing our results is based on the fact that the energy levels are not independent numbers to be chosen at random but that they form a closely related group. Their positions are fixed by the three moments of inertia together with a limited number of parameters which determine the centrifugal stretching correction. The circumstance that the levels of Table I obey these conditions, is shown by the regular behavior of the quantity δW. In most cases, the displacement of a level by as much as 1 or 2 cm^{-1} would introduce pronounced discontinuities in the series which may be formed from successive values of δW.

The essential consistency of the analysis is indicated by the excellent correlation between the observed and predicted spectrum as shown in Figs. 3 and 4. Every strong observed line has an accompanying strong predicted line. There is no single instance of a calculated line of measurable intensity which fails to appear in the experimental curve.[10] This, we regard as highly important. There are, on the other hand, six lines of medium intensity and seven weak lines which are not represented in the predicted spectrum. For the most part these lie towards the higher frequencies and undoubtedly are lines corresponding to transitions to the levels of $J=12$ or even 13. There would be no difficulty about postulating levels by means of which these 13 lines would be accounted for. This we have refrained from doing since from a few scattered levels of $J=12$ we could not observe those regularities in δW which have everywhere been considered essential to the analysis.[11]

The analysis of the rotational spectrum of

[10] The one possible exception lies at 249 cm^{-1} where two weak lines are predicted which should presumably have been observed. The reason that they were not, appears evident from a study of the original curves which show that great care was taken in drying the spectrometer as well as possible in order to measure accurately the two strong lines at 248.04 and 254.05 while less attention was paid to the intervening region.

[11] Certain levels of $J=12$ and higher are indeed listed in Table I but these were determined from lines which are members of a series and about which there can be little doubt.

water vapor appears to be complete in the sense that the lines have been identified and the energy levels determined. The next step is, of course, to obtain with great precision the effective moments of inertia of the molecule. This will demand an adequate calculation of the rotational stretching effect, a problem which we now plan to attack. On the experimental side we are endeavoring to map the rotational spectrum of D_2O. An accurate knowledge of the effective moments of inertia of both H_2O and D_2O would do much towards making possible a precise determination of the true moments and hence of the dimensions of the molecule.

AUGUST 1, 1937 PHYSICAL REVIEW VOLUME 52

Secondary Processes of Ionization in Mercury Vapor

BENJAMIN L. SNAVELY
*Princeton University, Princeton, New Jersey**

(Received April 16, 1937)

The ionization occurring in mercury vapor as the result of the impact of electrons of energy less than 10.4 volts was studied by use of a tube containing a small filament, the emission of which was limited by space-charge. The ionization produced has been studied as a function of electron velocity, bombarding current and pressure of the mercury vapor. For electron velocities above 7 volts the data are in agreement with the assumption that the larger part of the ionization is the result of ionization of metastable atoms by electron impact. Critical potentials at 8.4 and 8.6 volts are identified with the configurations $5d^96s^26p\ ^3P_1^0$ and $5d^{10}6s7p\ ^3P_1^0$, respectively. Another critical potential at 6.9 volts is tentatively identified as the difference between the lower metastable level, $6s6p\ ^3P_0^0$, and a negative energy level, $6p^2\ ^3P_1$. Large numbers of negative ions were found in the neighborhood of 4.9 volts. It is suggested that ionization observed in this region is due to the simultaneous formation of a positive ion and a negative ion upon the collision of a 2^3P_0 atom with a 2^3P_1 atom.

INTRODUCTION

A QUANTITATIVELY plausible explanation of the origin of the ionization which must occur in a low voltage arc was first given by K. T. Compton[1] in terms of an initial excitation of the vapor by electron impact and subsequent ionization of the excited atoms by the impact of other electrons. Various experiments in which the ionization is produced under more exactly controlled conditions than exist within an arc have since been explained on the basis outlined by Compton. Thus Franck and Einsporn[2] found critical potentials in mercury vapor corresponding to the potentials necessary to ionize the metastable levels and Smyth and Compton[3] observed a decrease in the ionization potential of iodine when it was excited by radiation from a mercury arc. Kannenstine[4] and Marshall[5] attempted to observe a lowering of the ionization potentials in He and in Hg due to the presence of metastable atoms but it was shown by Pool[6] that their results were inconclusive.

Experiments on the photoionization of mercury vapor by λ2537, discovered by Steubing[7] and later studied by Rouse and Giddings,[8] Foote,[9] and Houtermans,[10] indicate that there is probably some other secondary process in addition to that outlined by Compton which may be of major importance at least at higher pressures. It has been definitely shown that ions are produced as the result of two successive absorptions of λ2537 and that no other radiation is involved. Houtermans showed that the ions are produced probably

---

* Now at Lehigh University, Bethlehem, Pennsylvania.
[1] K. T. Compton, Phys. Rev. **15**, 476 (1920).
[2] J. Franck and E. Einsporn, Zeits. f. Physik **2**, 18 (1920).
[3] H. D. Smyth and K. T. Compton, Phys. Rev. **16**, 501 (1920).
[4] F. M. Kannenstine, Astrophys. J. **59**, 133 (1924).
[5] M. Marshall, Astrophys. J. **60**, 243 (1924).
[6] M. L. Pool, Phys. Rev. **30**, 848 (1927).
[7] W. Steubing, Physik. Zeits. **10**, 787 (1909).
[8] G. F. Rouse and G. W. Giddings, Nat. Acad. Sci. Proc. **11**, 514 (1925); **12**, 447 (1926).
[9] P. D. Foote, Phys. Rev. **29**, 609 (1927).
[10] F. G. Houtermans, Zeits. f. Physik **41**, 619 (1927).

as the result of a collision between a 2^3P_1 and a 2^3P_0 atom so that the total energy available is only 9.51 volts. This is 0.87 volt less than the energy required for ionization by a primary process. Houtermans suggested that in this case molecular ions are formed. Later Arnot and Milligan[11] found that Hg_2 ions could be produced by the impact of electrons having energies in excess of 9.5 volts.

Nielsen[12] found that negative ions were produced in mercury vapor by electron impact at 2.7, 4.9, 5.5, and 8.8 volts, probably as the result of some secondary process. Arnot and Milligan[13] also observed negative Hg ions due to the impingement of positive ions on a metal surface.

In the present paper the total ionization produced in mercury vapor by the impact of electrons having energies below 10.38 volts is studied with a view to identifying the processes responsible for the ionization.

The general experimental method consisted in studying the ionization produced in mercury vapor by bombardment with slow electrons, alone and in combination with the radiation from a low pressure mercury arc, as a function of the bombarding voltage, with various bombarding currents and pressures of vapor. Hertz's[14] method, in which the presence of positive ions increases the space charge limited current from a small filament, was used to determine the amount of ionization of the vapor.

The procedure was to obtain a curve showing the variation in the current i between the filament and the anode as a function of the accelerating voltage of the bombarding electrons, V. Ionization setting in at a given voltage would then be indicated by a change in the slope of the curve at the corresponding voltage. The curves obtained were analyzed quantitatively.

Apparatus

Figure 1 (a) shows the general type of tube employed. The tubes were made compact enough so that the complete assembly of anode, cathode and filament could be slipped down into the quartz portion of a 17 mm internal diameter quartz Pyrex graded seal which served as the envelope. This construction eliminated waxed or ground joints and made it possible to bring the mercury arc up quite close to the portion of the vapor within the tube which was to be illuminated. The oxide coated cathode was of the type described by Hertz and Kloppers.[15] The anode was of nickel, hydrogen fired and then degassed by means of an induction furnace after the tube had been pumped down and baked out. A part of one side of the anode was made of nickel gauze in order to permit the irradiation of the vapor within it. The filament was a 0.1 mm tungsten hairpin 8 mm long, generally run at a temperature estimated to be about 1700°C. Its potential was adjusted with respect to the anode so that the emission would be about 20 microamperes.

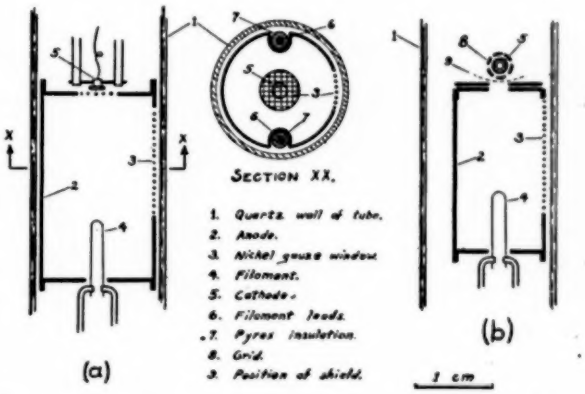


FIG. 1. Arrangement of electrodes in experimental tubes.

The tube shown in Fig. 1 (b) was designed originally to decrease any possible contamination of the anode by material evaporated from the cathode. It became apparent later that this effect was probably unimportant. However the curves obtained with this tube, due to the peculiarity of its construction, differed significantly from those obtained with the other tubes. A small shield of sheet molybdenum moved by an external magnet acting on a small iron armature shielded the anode from the cathode except when observations

[11] F. C. Arnot and J. C. Milligan, Roy. Soc. Proc. **153**, 359 (1936).
[12] W. M. Nielsen, Phys. Rev. **27**, 716 (1926); Nat. Acad. Sci. Proc. **16**, 721 (1930).
[13] F. C. Arnot and J. C. Milligan,' Roy. Soc. Proc. **156**, 538 (1936).
[14] G. Hertz, Zeits. f. Physik **18**, 307 (1923).
[15] G. Hertz and R. K. Kloppers, Zeits. f. Physik **31**, 436 (1925).

were actually being made. The cathode was a nickel cylinder about 4 mm long heated by an internal coil of tungsten. The tungsten spiral grid surrounding the cathode could, for cleaning purposes, be heated electrically. The magnitude of the bombarding current was fixed by the voltage applied between the cathode and the grid which was set for a given run and was between one and four volts accelerating. The variable bombarding voltage was applied between the cathode and the anode.

There were several tubes of the type shown in Fig. 1 (a), which were practically identical except for the shape and size of their anodes. In what follows they will be referred to as "tube *A*." The tube shown in Fig. 1 (b) will be referred to as "tube *B*."

In order to increase the mercury vapor pressure above that corresponding to room temperature the tube was enclosed within an oven and connected to the pumps through a U-tube mercury cut-off one arm of which passed through the bottom of the oven. The vapor was supplied from the mercury in this arm.

A device for recording the i–V curves directly gave considerably greater accuracy than was possible by any point by point method of recording and permitted a great decrease in the time required to obtain a curve. Essentially it consisted of a galvanometer placed in the circuit between the anode and the filament and a system of lenses and mirrors so connected to the slider of the potential divider supplying the bombarding potential that a beam of light would be deflected horizontally by changes in the filament to anode current i while the vertical deflection would be proportional to the bombarding voltage V. The beam was brought to a point focus on a photographic plate. In order to make the voltage variation as uniform as possible the slider on the potential divider was driven by an electric motor through a system of reducing gears. The speed was adjusted so that a curve would be completed in about 2 minutes. The galvanometer used in the recorder had a period of 2 seconds.

Bombarding potentials were read on a large laboratory standard voltmeter reading to 15 volts full scale. It was checked against a standard cell at the beginning of the work and again toward the end.

In using the recorder an axis corresponding to zero current was always placed on the plate by short-circuiting the recorder galvanometer and running the slider of the potential divider through its travel once. Voltage marks were placed either on the curves directly or on the current axis. This was done by turning off the lamp supplying the illumination for the recorder for an instant when the hand of the voltmeter crossed each volt mark.

A mercury arc with an oxide coated cathode was used as a source of λ2537. The arc discharge of about 2 amperes passed through a tube 1 cm in diameter, having a quartz central section. A side arm held at about room temperature kept the mercury vapor pressure within the arc at a fairly low value. In use the quartz portion of the arc projected up into the oven almost against and parallel to the experimental tube.

Results

Figure 2 is a typical i–V curve taken with tube *B* showing the principal features of the curves

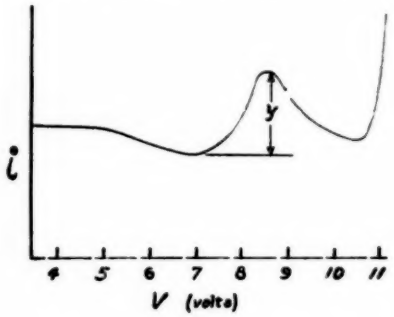


FIG. 2. Typical i–V curve made with tube *B*. $p=0.09$ mm. I varied between 50μa for $V=4$ volts, and 96μa for $V=10$ volts.

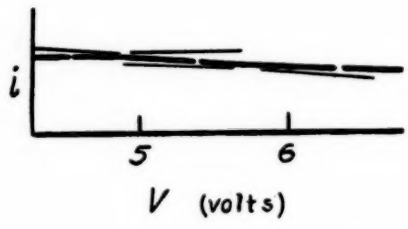


FIG. 3. Reproduction of portion of an i–V curve made with tube *A* showing breaks in the neighborhood of 5 volts.

studied. The voltage scale of each curve was corrected for contact potentials by translating it so that the beginning of the upward bend indicating primary ionization would occur at 10.4 volts.

Certain features of the curves may be noted:

(a) An upward rise beginning at 6.9 volts. For low vapor pressures this rise does not become noticeable until about 8 volts.

(b) A peak between 8 and 9 volts which from its shape appears to consist of two unresolved peaks at 8.3 and 8.6 volts.

(c) A decrease at 4.9 volts. This decrease was considerably larger for tube B than for tube A.

(d) A small upward break was observed on the curves taken with tube A which occurred about 0.8 volt higher than the downward break at 4.9 volts. Fig. 3 shows these breaks at 4.9 and 5.7 volts.

Figure 4 shows the effect of illumination from the quartz mercury arc on the filament-anode current. Both the curves shown here were taken under conditions which were as nearly as possible identical except that curve (a) was taken while the tube was illuminated by the arc and (b) was taken without such illumination. For purposes of comparison the upper curve has been transposed as shown by the dotted curve. The general shape indicates that the electron velocity distribution was considerably broader than in any of the curves used for locating the critical potentials. It is seen that as a result of illumination by the arc: (a) There is a general increase in the filament current. (b) The height of the peak between 7 and 10.4 volts is increased. (c) The increase in the height of this peak is relatively greater between 7 and 8 volts than at the higher voltages.

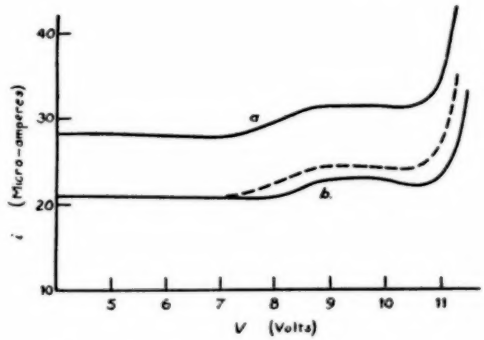


FIG. 4. Effect of radiation from the mercury arc on $i-V$ curves. These curves were traced from data recorded automatically with tube A. (a) Arc on. (b) Arc off. For the purpose of comparison curve (a) has been transposed as shown by the dotted curve. $I=60\mu a$. $p=0.03$ mm.

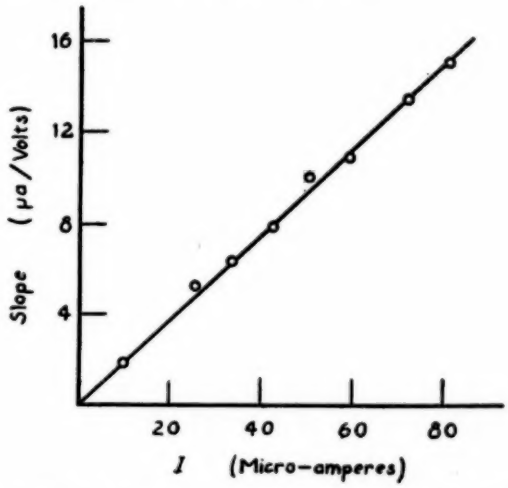


FIG. 5. Slope of $i-V$ curves at 11.2 volts as a function of the bombarding current in tube A. $p=0.23$ mm.

As a basis upon which to test the quantitative functioning of the filament as a detector of ionization, it was assumed that almost all the ions formed by electrons whose energies exceeded 10.38 volts would be produced as the result of a primary process. The slope of the $i-V$ curves at 11.2 volts was chosen as a measure of the change in the filament current due to the ionization. An inflection at this point reduces the effect of an error in choosing the point at which to measure the slope. In Fig. 5 this slope has been plotted against the bombarding current for constant mercury vapor pressure. The linear relation obtained is considered sufficient justification for the assumption that at constant pressure, changes in the current from the filament are directly proportional to the ionization.

In Fig. 6 the slope at 11.2 volts per unit bombarding current has been plotted as a function of the mercury vapor pressure, p, using logarithmic scales. Kingdon[16] and Foote and Mohler[17] have shown that the effectiveness of an ion in neutralizing space charge about a small filament varies inversely as the two-thirds power of the pressure. Under these conditions the effect of a given bombarding current on the filament

[16] K. H. Kingdon, Phys. Rev. **21**, 408 (1923).
[17] P. D. Foote and F. L. Mohler, Phys. Rev. **26**, 195 (1925).

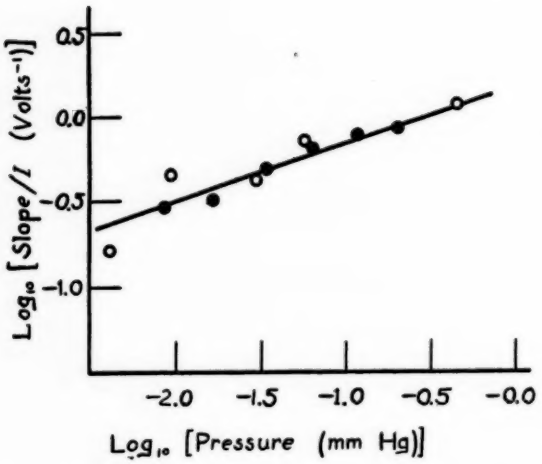


FIG. 6. Slope of $i-V$ curves per unit bombarding current as a function of the mercury vapor pressure. Points taken with two different tubes. The line through the points has a slope of one-third.

current should vary as $p/p^{\frac{1}{3}}=p^{\frac{1}{3}}$, where p is the vapor pressure. The proximity of the points of Fig. 6 to a line drawn with a slope of $\frac{1}{3}$ indicates that the points are in good agreement with this equation.

In the curve of Fig. 7 the height of the peak at 8.6 volts is shown as a function of the bombarding current I. The peak height was measured as indicated by y in Fig. 2. The curve in this illustration, however, has a minimum which is very much larger relative to the peak at 8.6 volts than any of the curves used for the data of Fig. 7 or Fig. 8. The points of these two curves would not have been appreciably changed if the peak height had been taken as the difference between filament currents at the peak and at $V=0$. The curve shown in Fig. 7 was derived from a certain hypothesis as to the origin of the peak which will be discussed later.

In Fig. 8 the peak height at 8.6 volts has been plotted as a function of the mercury pressure, the bombarding current being approximately constant.

The magnitude of the break at 4.9 volts was measured as the difference between the slopes of the $i-V$ curve at 4.5 and 5.3 volts. With tube A at a constant mercury vapor pressure and a bombarding current between 30 and 250 microamperes this break varied very nearly as the square of the bombarding current. In Fig. 9 the change of slope at 4.9 volts is shown as a function of the mercury vapor pressure for two different tubes. Because of considerable variations in the bombarding currents between the different points the changes of slope have been divided by the squares of the respective bombarding currents.

The change of slope at 5.7 volts was measured in the same manner as the break at 4.9 volts. This break was too small to permit any accurate measurements. For constant vapor pressure, it varied as some power between the square and the cube of the bombarding current for a fivefold variation in the latter.

The effect of nitrogen up to pressures of several mm was tried. The results which were somewhat erratic confirmed the conclusions obtained from mercury alone but showed nothing additional.

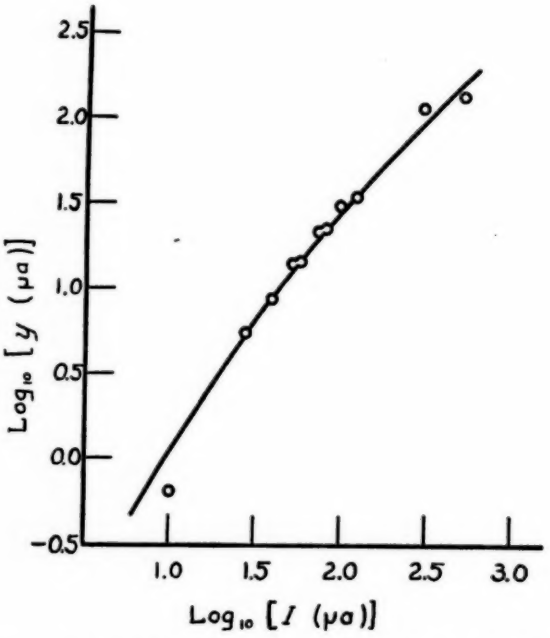


FIG. 7. Peak height at 8.6 volts as a function of the bombarding current. $p=0.23$ mm.

DISCUSSION

We may list the following as possible secondary processes by which ions might be formed:

(1) A collision of the first kind between an electron and an excited atom.

(2) A collision of two excited atoms.

(3) The absorption of a quantum by an excited atom. Unless the energy of the quantum corresponds very closely to the energy of an allowed transition from the excited state the

probability of such an absorption is very small and for the present discussion we shall assume it to be negligible.

We are here concerned chiefly with those excited levels which lie below 8 volts (Fig. 10). The lifetimes in these levels in seconds are of the order of 10^{-7} for 2^3P_1, 10^{-8} for 2^3S_1, 10^{-9} for 2^1P_1, and the lifetime of 2^1S_0 is probably within this region. These times are all considerably less than the mean free times calculated from kinetic theory which for the range of temperatures and pressures used here lie between about 5×10^{-5} and 7×10^{-7} seconds. It may be assumed therefore that the rate at which atoms disappear from these levels will be practically independent of pressure for the present range of data.

The undisturbed lifetimes of the metastable levels are very much greater, being of the order of seconds. Under laboratory conditions the lifetimes are reduced to the order of 10^{-2} to 10^{-5} seconds by the following processes: (a) Collisions of the first or second kind with other atoms. (b) Diffusion to the walls. (c) Absorption of quanta. (d) Collisions of the first or second kind with electrons.

The net effect of the first three of these processes may be estimated from certain experimental results obtained by Zemansky[18] who found that the decay constant, β, for the fluorescence of λ2537 in mercury vapor after the exciting radiation was removed could be expressed as:

$$\beta = aN + b/N,$$

where a is a constant, N is the atomic concentration which is approximately proportional to the pressure, and b is a constant depending on the geometry of the apparatus. On the basis of a theory advanced by Zemansky and of another developed by Samson,[19] β will also be the decay constant for 2^3P_0 atoms. In an absorption cell 1.95 cm thick used by Zemansky the two terms on the right side of the above equation were equal at a saturated mercury vapor pressure of about 0.2 mm corresponding to a temperature of the cell of 95°C. For a smaller cell, 1.3 cm thick, b was greater so that the two terms were equal at about 110°C, corresponding to a mercury pressure of about 0.5 mm. Since the internal dimensions of the tube used here were considerably smaller than the smallest absorption cell used by Zemansky, and since the pressure was always

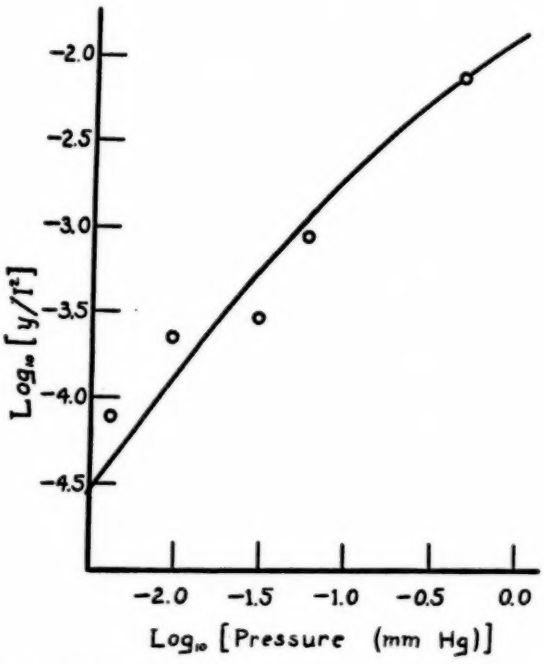


FIG. 8. Peak height at 8.6 volts as a function of the mercury vapor pressure. $I = 60\mu a$ approximately.

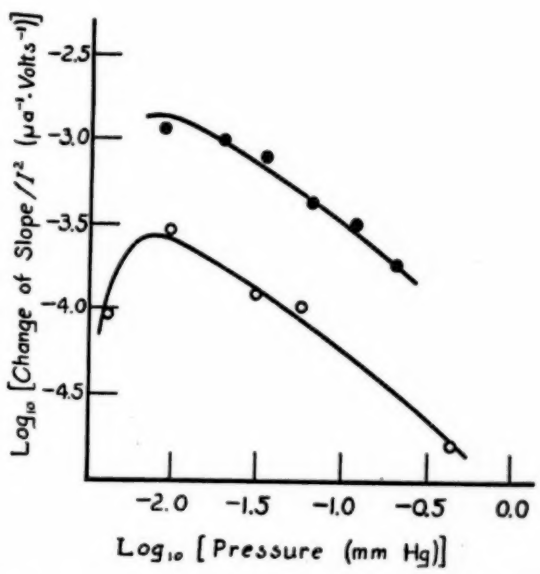


FIG. 9. Variation of the change of slope at 4.9 volts with mercury pressure.

[18] M. W. Zemansky, Phys. Rev. **34**, 213 (1929).
[19] E. W. Samson, Phys. Rev. **40**, 940 (1932).

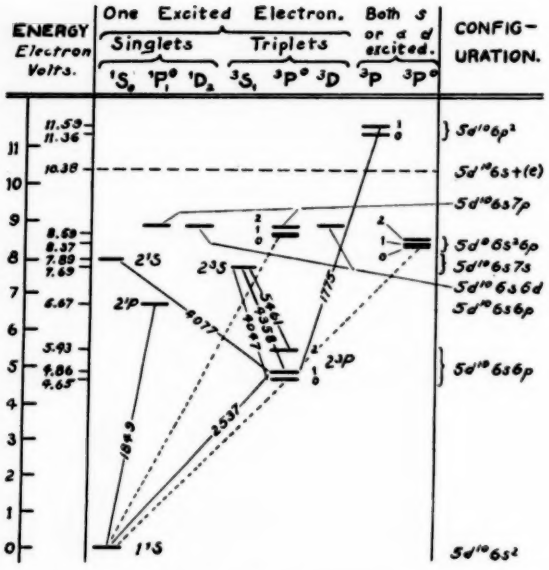


FIG. 10. Mercury energy levels referred to in this paper.

below 0.3 mm, it may be assumed, insofar as we neglect the effect of electronic collisions, that the rate of disappearance of 2^3P_0 atoms will be inversely proportional to the pressure.

The rate at which atoms are removed from the metastable levels by electronic collisions will be proportional to the bombarding current. Because of their comparatively longer lifetimes, this process may be of importance in the case of metastable atoms, while being comparatively negligible for ordinary excited states.

The general shape of the i–V curves in the neighborhood of 8 volts is very similar to the optical excitation functions of the lines λ4047, λ4358, λ5461, as determined by Siebertz.[20] These lines all originate on 2^3S_1. This suggests that the ionization occurring in the neighborhood of 8 volts is the result of an initial excitation to the 2^3S_1 level. If the energy of the bombarding electrons is increased beyond 7.69 volts where this level is first excited there will be a sudden increase in the number of atoms reaching the metastable levels because 2^3S_1 is the lowest level from which the metastable levels 2^3P_0 and 2^3P_2 may be populated by direct optical transitions. Since the difference in energy between a normal ionized atom and a metastable atom is only 5.74 volts if the latter is in the 2^3P_0 state and 4.95 volts if in the 2^3P_2 state, electrons having sufficient energy to excite the 2^3S_1 level will have more than sufficient energy to ionize atoms in either of the metastable levels. Furthermore since the lifetime of the metastable levels is such that an appreciable number of atoms in these levels may be struck by the bombarding electrons, it is reasonable to suppose that some ionization will result. The process may be treated quantitatively as follows:

Let Δi be the increase in the current from the filament to the anode as the result of this secondary process at some particular value of the bombarding voltage V.

Let n be the concentration of metastable atoms.

As before let I be the bombarding electron current.

Then since as we have already seen, the tube responds linearly to the rate of production of ions, and since the sensitivity varies inversely as the two-thirds power of the pressure,

$$\Delta i = kp^{-\frac{2}{3}}nI, \tag{1}$$

where k is a constant which depends on the tube.

If the metastable levels are supplied chiefly by radiation from the 2^3S_1 level, the rate of production will be directly proportional to the concentration of 2^3S_1 atoms. This concentration, insofar as we neglect all secondary methods of production, will be jointly proportional to the bombarding current and the concentration of normal atoms. If we neglect temperature variations, this last factor may be replaced by the pressure. Hence we have upon equating the rates of production and disappearance of metastable atoms:

$$aIp = bnI + cn/p, \tag{2}$$

where a, b, and c are constants which depend on the tube. The first term on the right side of this equation represents the rate at which electrons in the bombarding beam cause atoms to pass out of the metastable levels and the last term is the rate at which metastable atoms disappear due to other causes.[21]

[20] K. Siebertz, Zeits. f. Physik **68**, 505 (1931).

[21] Equation (2) gives the same relationship between the concentration of metastable atoms and the bombarding current as was found by Kopferman and Ladenburg, Zeits. f. Physik **48**, 15 (1928), to exist between the concentration of metastable atoms in the positive column of a neon glow discharge and the current density.

On combining Eq. (1) and Eq. (2) we have:

$$\Delta i = akI^2p^{\frac{1}{2}}/(bI+c/p). \quad (3)$$

Let y be the height of the peak at 8.6 volts. Then if the ionization is due almost entirely to the process outlined, we will have:

$$y = I^2p^{\frac{1}{2}}/(UI+V/p), \quad (4)$$

where U and V are constants which depend upon the tube.

Equation (4) may be written as:

$$I^2/y = (U/p^{\frac{1}{2}})I + V/p^{4/3}. \quad (5)$$

Hence for constant p, I^2/y plotted as a function of I should give a straight line of slope $U/p^{\frac{1}{2}}$ and intercept $V/p^{4/3}$. In Fig. 11 the data corresponding to Fig. 7 have been plotted in this manner. From the slope of the line drawn through these points, $U/p=3.0$; and from the intercept, $V/p=57$; where $p=0.23$ mm of mercury, and I is expressed in microamperes. These values have been substituted in Eq. (4) to compute the curve of Fig. 7.

Substituting for p its known value we obtain: $U=1.8$ and $V=8$. These are constants which are independent of current and pressure and depend only on the geometry of the tube and the operating conditions of the filament. The tube used for obtaining the points of Fig. 7 and the tube used for the points of Fig. 8 had filaments which were almost exactly alike, although the anodes differed in their dimensions. From the value of U and V determined by the former tube it was estimated, from a consideration of the differences in the anodes, that for the data of Fig. 8, we should have approximately, $U=0.9$ and $V=16$. These values were used in Eq. (4) to compute the curve shown in Fig. 8.

It appears reasonably certain that the peak cannot be due to either of the other secondary processes previously mentioned. If the peak were due to collisions of excited atoms with other excited atoms or with quanta, it would still be necessary to assume that metastable atoms were involved in order to explain the downward curvature of the points of Fig. 7, since these are the only states having lifetimes sufficiently long to make destruction by electron impact at all probable and thus result in a total ionization which varies by some power less than the square of the bombarding current. The concentration of the second constituent of such collisions, excited atoms or quanta, would vary approximately with at least the first power of the pressure because

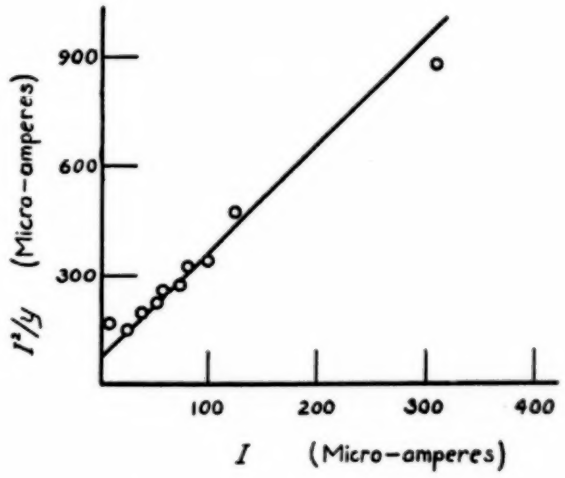


FIG. 11. I^2/y as a function of I.

the rate of production by electron impact depends upon the number of collisions which the bombarding electrons make with normal atoms. With such a mechanism Eq. (4) would be replaced at sufficiently low pressures or small bombarding currents, by an equation of the form:

$$y = (1/V)I^2p^n,$$

where $n \geqq 7/3$, and the points of Fig. 8 might be expected to approximate to a line of slope 7/3. The slope of approximately one which is actually observed agrees much better with the original hypothesis as to the origin of the peak.

It seems very probable that of the two metastable levels, 2^3P_0 must be the one chiefly involved in the processes of ionization discussed so far. The term in Eq. (2) which we attributed to destruction of metastable atoms by collisions with electrons implies a metastable state having a fairly long life. This condition is fulfilled much better by the 2^3P_0 state than by the 2^3P_2. It is principally the former level which is filled by radiation from the mercury arc so that the additional ionization produced is most reasonably ascribed to the excitation of this state rather than to 2^3P_2.

It is probable that the sharp maximum occurring in the excitation function of the level

2^3S_1 and the sharp maximum also observed here in the secondary ionization may be due to the fact that when the energy of the electrons equals or exceeds that necessary to excite the levels just above 2^3S_1, the number of atoms reaching 2^3S_1 suddenly decreases because the energy of the bombarding electrons is now taken up by the excitation to the higher levels. This hypothesis is in beautiful agreement with the shape of the peak obtained here between 8 and 9 volts. The first two levels above 2^3S_1 to which transitions from the ground state are optically allowed, and hence which might be expected to have a high probability of excitation under electron impact are $5d^9 6s^2 6p\ ^3P_1^0$ and $5d^{10} 6s 7p\ ^3P_1^0$ at 8.37 and 8.59 volts respectively. These voltages correspond very closely with the sharp downward breaks obtained here at 8.3 and 8.6 volts. Since these breaks occur on a portion of the curve which is rising very rapidly they give the appearance of peaks. The voltage difference between these peaks as determined from the six best $i-V$ curves obtained with tube A is 0.25 ± 0.04 volts, which is in agreement with the difference of 0.22 volt between the two $^3P_1$ terms.

If we accept as correct the foregoing interpretation of the peaks between 8 and 9 volts, their positions may be used to correct for contact potentials more accurately than is possible with the somewhat rounded break indicating ionization at 10.38 volts. On this basis the upward rise near 7 volts is found to lie at 6.90 ± 0.05 volts. This value was the average of six determinations taken from the same plates as were used to calculate the voltage differences between the peaks.

It seems reasonable to attribute the break at 6.9 volts to an ultraionization potential of the lower metastable state. The position of this break corresponds closely with the energy of 6.93 electron volts required for the transition, $6s6p\ ^3P_0^0 \rightarrow 6p^2\ ^3P_1$. $6p^2\ ^3P_1$ is a negative energy level having a term value of -9789 cm^{-1}. This is an optically allowed transition and therefore might be expected to have a fair probability of occurrence under electron impact. Shenstone[22] has suggested that ultraionization potentials may be excitation potentials of negative energy levels which produce ions by autoionization. However, as Shenstone has pointed out, it is not to be expected that the $6p^2\ ^3P$ levels would lead to such ionization directly. Ionization would have to be preceded by a radiative transition to a lower unstable level.

The break at 5.7 volts is probably an ionization potential of the 2^3P_0 level. This agrees with the difference of 5.73 volts between the ionization potential at 10.38 volts and the excitation potential of the 2^3P_0 level.

The origin of the downward break at 4.9 volts is not entirely clear. Off-hand one might expect 4.9 volt electrons to have about the same effect on the ionization within the tube as the absorption of λ2537, which has the same energy. λ2537 excites the 2^3P_1 level and from this the 2^3P_0 level is populated by collisions with normal atoms. These are the only two levels which we would expect to be primarily excited by 4.9 volt electrons. However, it is observed that where absorption of λ2537 leads to an *increase* in the filament-anode current, the impact of 4.9 volt electrons leads to a *decrease* in the filament-anode current.

Since the magnitude of the break at 4.9 volts varies as the square of the bombarding current, it seems improbable that the effects observed can be due directly to any space charge effects of the bombarding electrons. For the same reason and also because of the reproducibility of the break with different tubes it appears very improbable that it is due to the accumulation of charges on any insulating layers which might have formed inside the anode. Electrons striking the filament and its supports might cause a slight decrease in the apparent current between the filament and anode as measured by an external galvanometer, but there would be no simple explanation as to why the effect is observed only above 4.9 volts. We are left with the conclusion that the downward break at 4.9 volts probably indicates the presence of negative ions.

Although there is not sufficient evidence to state definitely the mechanism which causes the downward break at 4.9 volts, the following hypothesis is in accord with known facts. Houtermans[10] has shown that the collision of a 2^3P_1 and a 2^3P_0 atom is probably responsible for the ionization accompanying the absorption of

[22] A. G. Shenstone, Phys. Rev. **38**, 873 (1931).

λ2537. Suppose that the collision of these two atoms results in the transfer of an electron so that a positive and a negative ion are formed simultaneously. The minimum total energy necessary for this process would be equal to the difference between the ionization potential and the electron affinity of the normal atom. Since the total energy available in the two excited atoms is $4.65+4.86=9.51$ volts, and since the ionization potential is 10.38 volts, the reaction suggested above will be energetically possible provided the electron affinity is in excess of $10.38-9.51=0.87$ volts. Glockler[23] using an empirical extrapolation method has estimated the electron affinity to be about 1.79 volts.

We would expect on this basis that, where the excitation is produced by radiation, the filament current would be increased because the negative space charge around the filament would cause diffusion of positive ions into that region. On the other hand where excitation is produced by electron impact there will always be a large number of slow electrons which by recombination with positive ions and by their space charge effects in other parts of the tube may cause the diffusion of an excess of negative ions into the region of the filament. This hypothesis is borne out by the results obtained with tube *B* in which the decrease at 4.9 volts was much more pronounced than in tube *A*. In the former tube the bombarding electrons were shot into the anode through a fairly large opening which permitted some penetration of the field from the grid into the space where ions were being formed. This field would act to pull positive ions out of the anode and to accelerate negative ions into it thus producing a greater preponderance of negative ions in the vicinity of the filament.

The decrease in the magnitude of the break at 4.9 volts as the pressure is raised is probably due to a decreased effectiveness of the stray fields within the anode in separating the positive and negative ions at the higher pressures rather than to a decrease in the actual amount of the ionization.

It seems very improbable that the downward break at 4.9 volts can be due to the process discovered by Arnot and Milligan[13] in which negative ions are formed when positive ions impinge on a metal surface. The probability found for this process decreased with the speed of the positive ions and was only 1.7×10^{-5} for 10 volt ions. Thus in any tube of the type used here the number of positive ions necessary for the production of the negative ions would completely prevent the detection of the latter. Furthermore if we did accept this explanation of the downward break at 4.9 volts, it would then be necessary to explain why the negative ionization did not mask the presence of positive ions in the region above 7 volts.

The writer takes this opportunity to express his deep gratitude to Professor L. A. Turner for the discussions which led to the problem of the present paper and for his friendly and indispensable criticism and advice during the course of the work. To Princeton University the writer is indebted for a fellowship held during the earlier stages of the work. It is also a pleasure to acknowledge to the physics department of Princeton University, appreciation for their indulgence in extending the facilities of Palmer Laboratory for several years following the termination of the author's residence in the university.

[23] G. Glockler, Phys. Rev. **46**, 111 (1933).

Carrier Mobility Spectra of Spray Electrified Liquids

SEVILLE CHAPMAN
Department of Physics, University of California, Berkeley, California
(Received March 15, 1937)

The mobility spectrum of spray electrified salt solutions has been investigated with an Erikson mobility tube under conditions of high resolving power. In general there are no charged carriers of mobility greater than 1.7 cm/sec. per volt/cm. At about this value the curves rise sharply, and then level out, approaching a broad maximum in the region between 0.05 and 0.10 cm/sec. per volt/cm, and then very gradually decrease toward zero. Superposed on this "background" striking peaks are observed, indicating groups of unique mobility. For the spraying of distilled water the most prominent peaks occur at mobilities of 1.5 for the negatives and 0.9 for the positives. Carriers of both signs are produced in about equal quantities. For bubbling the negative predominate in the ratio of 2.5 to one, and the prominent peaks occur at 1.5 and 0.3 for the negatives, and 0.9 and 0.4 for the positives. The introduction of salts to the water increases the electrification from spraying up to a concentration of about 1.0×10^{-4} normal. At this value the electrification is about double that for distilled water, but at higher concentrations the electrification decreases so that for 0.2 normal it is only 10 percent of that for distilled water. The salt gives rise to very strong peaks at mobility 0.5. Investigations have also been made of the effects on the curves of age of the carriers and humidity of the air. It is concluded that the peaks represent stable groupings, the negative 1.5 group being due to a normal negative ion in moist air coming from an evaporated drop. The 0.5 peaks in the salt curve presumably come from carriers containing electrolytic ions. Probably all of the carriers are singly charged. It is impossible to calculate quantitatively the size of the small carriers in the faster groups from their mobility, but from the value of the potential difference of the double layer, and estimates on salt solutions, it is concluded that the separation of the double layer is 2.0×10^{-6} cm corresponding to an extra electronic charge for every 550,000 H_2O molecules.

THE electrification of liquid droplets produced by spraying, bubbling, and similar methods is well known.[1–8] These phenomena depend upon the surface condition of the liquid, the presence of an electrical double layer within the surface being the important factor.[4]

While water and aqueous solutions have been studied by various workers (particularly by Lenard), the nonaqueous liquids have received relatively little attention. Results obtained by the writer[8] who used the Millikan oil-drop method on organic liquid droplets of 3×10^{-5} to 5×10^{-4} cm radius were quite different from those of Lenard, who estimated that his droplets were of 10^{-6} to 10^{-7} cm radius. To correlate the two sets of results, the writer set up an Erikson mobility tube capable of covering a wide range of mobilities under conditions of *high resolving power*. The present paper reports the results obtained with water and salt solutions. A study of nonaqueous liquids is now in progress.

Apparatus and Procedure

The apparatus is shown schematically in Fig. 1, which is, however, not drawn to scale. The spray is generated by a glass atomizer S, in the spray chamber M. The spray is led through the inlet I, into the Erikson tube. A blast of air from the air pressure regulator A_2 blows the spray down the tube, while the electric field between the plates P_1 and P_2 drives the charged carriers across the tube. The point at which they strike the plate P_2 will thus depend on their mobility. Hence by moving the electrodes E_1 and E_2 up or down the tube one can obtain a mobility spectrum, that is, a curve of current to the electrometer Q, as a function of electrode position.

The Erikson tube, made of poplar and coated on the inside with paraffin, has an over-all length of 3 meters. The air blast travels down the tube 80 cm before reaching the inlet I. The brass plates P_1 and P_2 are 140 and 200 cm long, respectively. The grounded plate (P_2) may be moved vertically over a range of 50 cm. Thus with the two electrodes E_1 and E_2, spaced 50 cm

[1] Lenard, Wied. Ann. **46**, 584 (1892).
[2] J. J. Thomson, *Discharge of Electricity Through Gases* (1898), p. 17 *et seq.*
[3] Coehn and Mozer, Ann. d. Physik **43**, 1048 (1914).
[4] Lenard, Ann. d. Physik **47**, 463 (1915).
[5] Blackwood, Phys. Rev. **16**, 85 (1920).
[6] Busse, Ann. d. Physik **76**, 493 (1925).
[7] Bühl, Kolloid Zeits. **59**, 346 (1932). A review.
[8] Chapman, Physics **5**, 150 (1934).

apart, it is possible to cover a range of 100 cm in downstream distance from the spray inlet I. P_1 supported on hard rubber is spaced 1 mm from the walls of the tube. The spacing between P_1 and P_2 is 2.78 cm. The air blast thus travels down the rectangular area 6.10×2.78 cm.

The spray chamber M is a virtually air tight brass box $13 \times 25 \times 40$ cm. The spray is led through the brass inlet I, 0.5×4.0 cm in cross section and 6 cm long. At the entrance to the tube, the inlet is bent over as shown, allowing an entrance of 4.0×0.25 cm. The inlet makes electrical connection between P_1 and M.

The solution to be studied is placed in the supply vessel G, and is siphoned into the reservoir J, which maintains constant level. The sprayer S and all the supply system is constructed so that, except for the platinum electrode D, the solution comes in contact only with glass. The top of the glass sprayer capillary tube is shielded on the outside with platinum foil connected to D. The sprayer is placed so that at least a large part of the spray generated goes into the inlet I.

The electrodes E_1 and E_2 are 0.5×4.0 cm with an air space 1 mm wide between them and the grounded plate P_2. The electrodes are supported by the sulphur block H, and protected by the spray guard K. A stream of filtered air dried over P_2O_5 comes from the filter N, and ventilates the electrode. This electrode ventilation was particularly necessary when working with concentrated salt solutions. It might be supposed

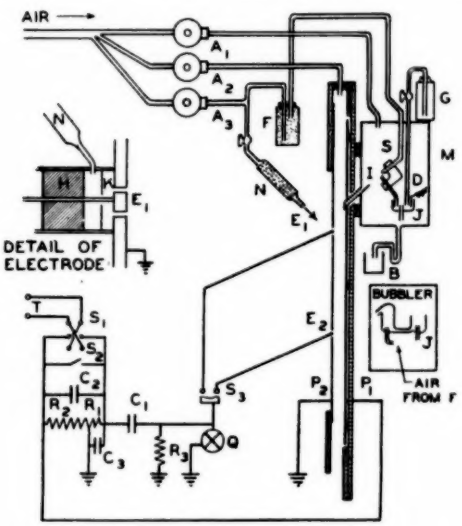


FIG. 1. Diagram of Erikson mobility apparatus.

that the ventilating stream would blow the charged carriers away from the electrode, but this is not so, as is proved by the fact that curves for weak solutions made with and without ventilation are found to be identical.

The air pressure regulators will be described elsewhere. The pressures of A_1 and A_2 are held constant within 1 percent, that of A_3 to $\frac{1}{2}$ percent. Throughout this experiment the air through A_2

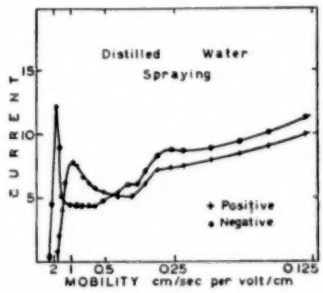


FIG. 2. Mobility spectrum of spray electrified distilled water.

was held at 2900 cm³ per sec., the value being obtained by use of a gas meter and water manometer. The air supply to the spray chamber M was 200 cm³/sec. This auxiliary supply doubled the number of carriers in the main blast. The air to the sprayer (60 cm³/sec., pressure equivalent to 19 cm of Hg) was filtered through the glass wool and cotton filter F. The average velocity of the air down the tube was 186 cm/sec. No turbulence was observed through a glass window when smoke was sent down the apparatus.

The electrometer (sensitivity 400 mm/volt) was shunted with a resistance of 3×10^{10} ohms so that the currents of 10^{-11} amp. could be easily measured. A transformer and rectifier supply the high voltage to the terminals T. A compensating circuit is introduced to permit a continuous variation of the voltage from zero to ± 4600 volts without inducing disturbing charges on the electrometer.

The mobility of the ions is given by the equation

$$k = dT/Vwl$$

where k is the mobility of carriers in cm/sec. per volt/cm; d is the (average) distance that the ions travel across the tube; T is the volume of air in cm³/sec.; w is the width of the tube; l is the distance downstream to the collecting electrode;

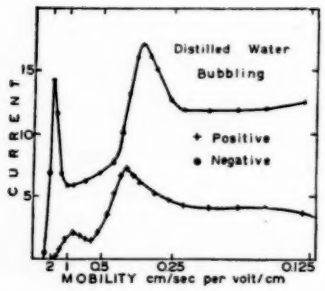


FIG. 3. Mobility spectrum by bubbling distilled water.

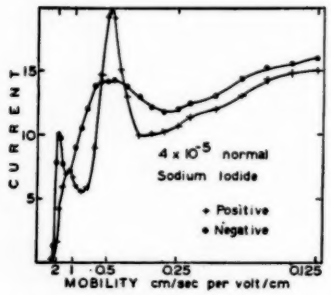


FIG. 4. Mobility spectrum of spray electrified sodium iodide solution.

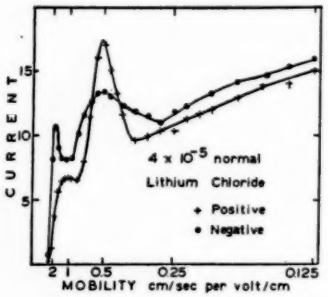


FIG. 5. Mobility spectrum of spray electrified lithium chloride solution.

and V is the voltage difference between the plates. It is possible to vary T, l, or V but in each run V only was varied. The usual effective downstream distance was 2.5 cm but measurements were made from 1.0 to 100 cm, which range permitted measurements of mobilities above 0.003 cm/sec. per volt/cm. Corrections as determined by auxiliary experiments with a model in a salt water trough were made for the distortion of the field near the inlet.

The distance of 2.5 cm downstream was chosen as the optimum value because for longer distances the sharp peaks disappeared due to diffusion, attachment of impurities, etc. Shorter distances gave a reduced resolving power. The time spent in the tube was thus $2.5/186 = 1/75$ sec. and the total time for the collection of the carriers after their production was certainly less than 0.2 sec., possibly less than 0.04 sec. Blackwood[5] mentions no time less than 2 sec.; dimensions of Bühl's[7] apparatus indicate times of at least 1 sec. The short time interval has an important bearing on the results obtained.

EXPERIMENTAL RESULTS

The experimental results are given in Figs. 2–14. The ordinates are given in the same arbitrary units. The abscissae are indicated in terms of mobility although the voltage is the factor determined directly in the experiment. While an air-blast method gives good relative values of mobility, the absolute magnitude may be in error by as much as 20 percent.[9] The curves rise sharply at a mobility of about 1.7 cm/sec. per volt/cm and then level out with some peaks superposed on this general background.

Spraying

Figure 2 shows the type of curve obtained for ordinary fresh distilled water. The striking feature of the curve is the strong negative peak at mobility 1.5 cm/sec. per volt/cm. The high mobility positive peak is not so strong, nor so sharp, and comes at mobility 0.92. There are weaker negative peaks at 0.39 (very weak, but definite as shown by duplicate runs) and 0.25; and a weaker positive peak at 0.29. Except in the immediate vicinity of the peaks the amounts of positive and negative electrification are about the same.

Bubbling

Figure 3 shows a curve made the same day by bubbling. In bubbling the breaking up of the surface is much less extensive in depth than in spraying since the film of water comprising a bubble is less than a wave-length of light in thickness just before bursting. Accordingly there should be a preponderance of negatives, since these come from the outer surface of the double layer. This is seen to be the case, the ratio of negatives to positives being about 2.5 to one. Variation in the bubbling pressure by a factor

[9] *Footnote added in proof:* Recalibration of static voltmeter indicates that all mobilities should be increased by 7 percent.

of four did not change this ratio, nor does the one to one ratio for spraying change over a limited (factor of 3) pressure range. Bühl[7] has observed a variation, which may have been due to working in a different pressure range, or to the fact that his carriers were of greater age.

With bubbling the same high mobility peaks occur as with spraying, but the lower mobility peaks are much more prominent in the case of bubbling. However, with very low spraying pressures the lower mobility peaks become more prominent. Fig. 3 is the only curve made by bubbling, all others were made by spraying.

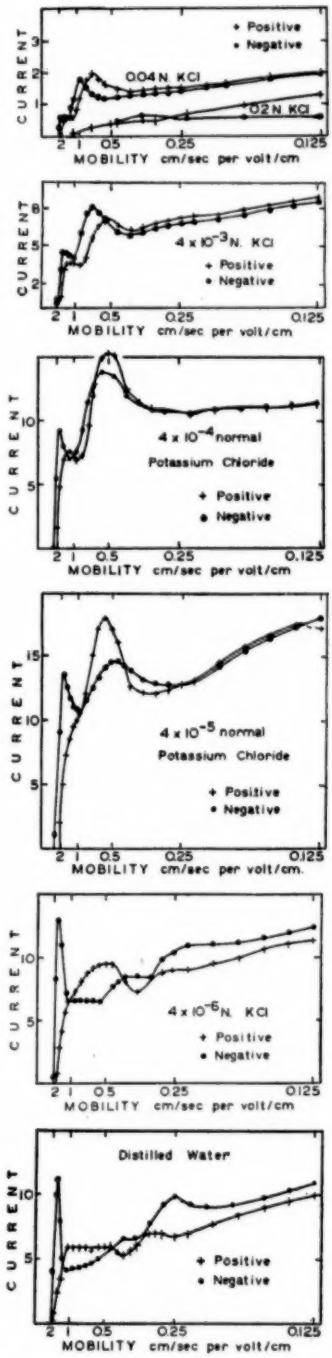
Salt

Figure 4 shows a typical curve obtained for salt solutions, in this case the solution being 4×10^{-5} normal sodium iodide. In addition to the high mobility peaks previously observed by bubbling and spraying distilled water, there are very prominent peaks at mobilities of 0.5. Fig. 5 for LiCl is almost identical with that for NaI, showing that the nature and size of the electrolytic ion are not of great importance in determining the mobility of the carrier (unless these electrolytic ions are hydrated to the same size).

Owing to the finite size of the inlet and the electrode, a group of unique mobility will be made evident by a peak of finite width on the mobility curve. One may calculate the exact shape of such a peak. In practice the "corners" of the curve are rounded somewhat, but with this exception the theoretical curve for the LiCl positive peak at mobility 0.52 coincides with the experimental curve within the accuracy of drawing the curves. Similar remarks apply to most of the very strong peaks observed, particularly those of mobility 0.5 for the positives, and 1.5 for the negatives. The striking sharpness of the strong peaks confirms the smoke test for the absence of turbulence in the Erikson tube.

Change of concentration

Figures 6 to 11 show the curves obtained for KCl for various strengths of solution. Fig. 6 shows the curves for 0.2 and 0.04 normal. No peaks are seen for the 0.2 normal, and the amounts of electrification are very small. For 0.04 normal there is some definite structure to



FIGS. 6–11. Mobility spectra of spray electrified potassium chloride solutions with varied concentrations.

the curves. With decreasing concentration the peaks increase in strength, and the amount of electrification increases until a strength of about 10^{-4} normal is reached. With further dilution the curves approach those for distilled water,

shown in Fig. 11. It is interesting to notice how the mobility of the peaks changes with concentration, particularly in the case of the 0.5

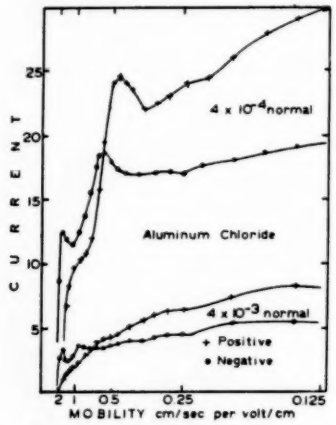


FIG. 12. Mobility spectrum of spray electrified aluminum chloride solutions.

negative group which changes from 0.4 for weak solutions to 0.9 for strong ones. Fig. 12, for $AlCl_3$, is similar to the other salt curves. Even HCl and NaOH give curves similar to those for salt.

The fact that the amount of electrification from salt solutions decreases with increasing concentration (above 10^{-4} normal) makes it impossible to account for much of the ionization observed over the oceans by a spray electrification process, since sea water is about 0.7 normal and even with 0.2 normal KCl the electrification is very slight.

Effect of age

Figure 13, for $MgSO_4$, is also similar to the other salt curves. The dotted curve is the curve for the negative carriers when the electrode was downstream 5.5 cm instead of the usual 2.5 cm. Thus the carriers had an additional 0.016 second to age. Even in this short time the high mobility carriers have become definitely less numerous, although the position of the peaks has remained unchanged. The positives behave similarly to the negatives, although to include the positive curve in Fig. 13 would be confusing.

Figure 14, for distilled water, was made by pointing the sprayer directly away from the inlet *I*, so that it was a matter of seconds before the carriers reached the Erikson tube. Comparison with Fig. 2 shows that most of the high mobility particles have disappeared due to the longer time for aging in moist air. It is not surprising therefore that previous workers who worked over long time intervals of several seconds should have failed to find high mobility carriers.

Interpretation of the Results

Distilled water peaks

The carriers comprising the negative 1.5 mobility peak have come from droplets which have evaporated down to nearly zero radius, since they are measured in relatively dry air. There is left one extra electron with a few water molecules or molecules from the air attached to it. Thus one has a "normal" negative ion in moist air. The same mechanism accounts for the high mobility positive peaks, except that there is less chance of spraying a *small* positive droplet since the outer surface of the double layer is negative. Thus the magnitude of the initial positive peak is less than that of the negative.

Salt peaks

The definite appearance of the positive and negative groups of mobility 0.5 with the addition of salts to the water is difficult to explain. These

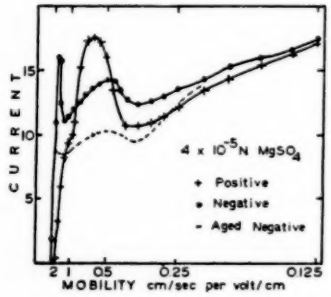


FIG. 13. Mobility spectrum of spray electrified magnesium sulphate solution. The dotted curve represents an aged curve for the negative carriers (see text).

groups may be "ionic" rather than "electronic." Apparently the presence of a Cl^- ion instead of an electron inside a few water molecules changes the mobility from 1.5 to 0.5. It is known that ions are hydrated in solution, although the exact nature of an hydrated ion is not certain. It may be that these experiments offer evidence for certain stable groupings of molecules about ions in solutions.

The cations such as K^+ will migrate into the outer negatively charged layer of the double

layer, leaving the anions such as Cl^- behind in the inner layer. Thus one would expect the $+0.5$ peak to be higher and narrower than the -0.5 peak in the same way that the -1.5 peak is higher and narrower than the $+1.0$ peak, since the higher and narrower peaks will come from the outer layer. Reference to a typical salt curve such as Fig. 4 for NaI confirms these conclusions.

Bühl[7] and others have pointed out that with increasing concentration the greater numbers of electrolytic ions present tend to neutralize the double layer so that it becomes more difficult to get charged carriers. This explains the decrease in electrification with increasing concentration of salt. The shift in the mobility of the -0.5 peak to higher values with increasing concentration may be due either to this neutralization or to the relative decrease in the number of water molecules available to each K^+ ion, thus decreasing the size of the carrier and increasing the mobility of the group.

Other peaks

In the past it has been customary to calculate the radius of the carrier from its mobility using one of the many mobility equations.[10] It is true that there are a number of mobility equations of various degrees of reliability and applicability, but unfortunately there is at the present time no rigorously derived single equation which covers a wide range of mobilities.[11] There are equations deduced for point charges, in which the dielectric forces between the ions and the molecules in the gas are predominant in determining the mobility. These agree fairly well with experiment for small ions (of high mobility). There are "solid elastic sphere" equations (which neglect the effect on the surrounding molecules of the charge on the ion) in which the physical radius of the ion is predominant. This type of equation is applicable to large particles (say of radius 10^{-6} cm or larger). For water, which is particularly polarizable, the region in which it is desired to obtain the size of the carriers (mobility 0.1 or greater) is just that where the two effects are both important. Apparently the most reliable equation which is in a form in which it can be readily applied, and which takes into account both of these effects, is the theory of Langevin.[12] In this equation, however, the size of the carrier enters in such a way that the mobility is relatively insensitive to variations of size in the region where the radius is about 10^{-7} cm. As a result it is not possible to arrive at a *quantitatively* accurate knowledge of the size of the carriers whose mobility may be of the order of 0.5 or less. Consequently the nature of the other peaks observed cannot be definitely established beyond the fact that they too represent more or less stable groupings.

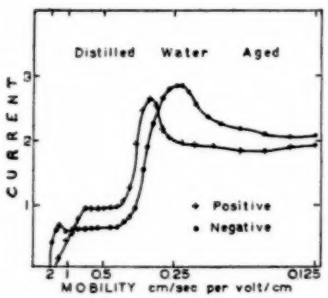


FIG. 14. Aged mobility spectrum of spray electrified distilled water.

Double layer

From the size of the carriers deduced from mobility measurements Lenard (see discussion on page 512, reference 4) has concluded that the negative layer extends to 80×10^{-8} cm and the positive layer to 150×10^{-8} cm below the surface, corresponding to a separation of 75×10^{-8} cm. Bühl quotes Lenard as saying that one extra electron per 50,000 molecules will account for the electrification.

Another method of calculating the separation is as follows. A careful inspection of the curves shows that the maximum height for the 0.5 peaks for KCl comes at a concentration of about 1.0×10^{-4} normal, and that the mean ordinates at this concentration are about double those for distilled water. It seems reasonable to suppose that at this concentration the number of ions from the salt correspond to the number of extra electrons associated with the double layer in pure water. Then using the value of the potential

[10] For instance Bühl, reference 7, refers to Lenard's equation. Lenard, Ann. d. Physik **61**, 665 (1920).

[11] See Loeb, *Kinetic Theory of Gases*, second edition, Ch. XI where the subject is discussed in detail.

[12] Ann. Chim. Phys. **8**, 238 (1905). See also Hassé, Phil. Mag. **1**, 139 (1926).

difference of the double layer as 0.0055 volt as determined experimentally by Bühl,[7] one may calculate the separation of the double layer.

The energy of thermal agitation at room temperature (about 0.04 volt) is so much larger than the potential difference of the double layer, that one would not expect a large permanent excess of salt in the double layer due to electrostatic attractions. The Gibbs adsorption has been investigated experimentally by McBain and Swain[13] who find that the adsorption for salt solutions amounts to a surface deficiency of salt corresponding to the positive adsorption of only one or two molecular layers of water. Thus one may conclude that the concentration of salt in the double layer is about the same as in the body of the liquid.

Accordingly on this basis the surface region contains one extra electron for every 550,000 H_2O molecules, and the separation of the double layer is 2.0×10^{-6} cm. These values, while only very approximate, do not involve the calculation of the size of any carriers, and hence are probably more accurate than estimates based on mobilities.

Effects of humidity and age

Previous workers in this field have generally blown the carriers down their mobility tube by the same blast of air which generated the spray. Thus their humidity may have been 100 percent. In this work the humidity of the main blast was about 50 percent, a significant difference. When all *three* blasts were simultaneously humidified, difficulties were encountered with the insulation so that complete curves could not be obtained. However, several points were obtained. In particular, with 100 percent humidity (taking account of leakage) the negative 1.5 peak dropped to less than 10 percent of its former value. Thus it is clear that the appearance of high mobility particles is due to evaporation of small unstable droplets.

The disappearance of the high mobility particles (see Fig. 14 for distilled water aged) is due to recombination, diffusion, or attachment to slower carriers. Blackwood[5] has shown that for the slower carriers of several seconds age, the shift of the maximum in the mobility curve (in these experiments, the broad maximum of the "background" at mobility 0.05 to 0.10) can be ascribed to a growth of the carriers at constant rate. Extrapolating from Blackwood's data, mobility 0.10 corresponds to an age of one second. Thus it is likely that the broad maximum observed in these experiments is due to particles which drifted in the spray chamber for about a second before entering the Erikson tube. The high mobility particles come from droplets sprayed directly into the inlet whose total age is 0.1 second or less.

There are four reasons why these groups of carriers have not been observed previously. Other work has been characterized by (1) lack of adequate resolving power, (2) moist air inhibiting evaporation, (3) ages of one second or more, and (4) in the case of salt solutions concentrations of 10^{-3} normal[6] or greater, while the concentration where the very strong peaks occur is an order of magnitude weaker.

Charge on carriers

Chapman[8] found with the Millikan method that the magnitude of the charge for the large carriers decreased with decreasing radius so that for drops of radius 10^{-6} cm or smaller, one would expect not more than one charge. In addition if the small carriers of fairly high mobility had many charges originally, they would tend to recombine readily, so that after 0.01 second or so, there would be few of them left. Thus it is likely that the carriers studied in this investigation are singly charged. The divalent and trivalent ions of $MgSO_4$ and $AlCl_3$ give the same typical curves as KCl. Unless the charge on very small droplets does not affect their mobility (see Langevin equation[12]), then the 0.5 carriers containing the Al^{+++} ion are still only singly charged (due to two extra electrons from the double layer). The preponderance of the positives for the slower $AlCl_3$ carriers is due to the three charges on each aluminum ion penetrating into the outer layer of the double layer.

In conclusion the writer wishes to thank Professor Leonard B. Loeb, who suggested the research, for his very helpful counsel and assistance.

[13] Proc. Roy. Soc. **A154**, 608 (1936). See also Harkins and McLaughlin, J. Am. Chem. Soc. **47**, 2083 (1925).

The Structure of Electronic Excitation Levels in Insulating Crystals

GREGORY H. WANNIER
*Princeton University, Princeton, New Jersey**

(Received May 13, 1937)

In this article, a method is devised to study the energy spectrum for an excited electron configuration in an ideal crystal. The configuration studied consists of a single excited electron taken out of a full band of N electrons. The multiplicity of the state is N^2. It is shown that because of the Coulomb attraction between the electron and its hole $N^{2/3}$ states are split off from the bottom of the excited Bloch band; for these states the electron cannot escape its hole completely. The analogy of these levels to the spectrum of an atom or molecule is worked out quantitatively. The bottom of the Bloch band appears as "ionization potential" and the Bloch band itself as the continuum above this threshold energy.

FOR several years, there have been two competing pictures in use to describe the behavior of electrons in crystals. The one adopted in most theoretical calculations and especially successful for metals describes each electron by a running wave, but Frenkel has shown that in many cases the more elementary atomic picture may be the better approximation.[1] This apparent contradiction has been removed by Slater and Shockley,[2] who showed with a simplified model that the two types of states actually coexist in a crystal. It is the purpose of the present paper to treat this question in a quantitative way, starting out from the actual Hamiltonian of the system.

We shall restrict ourselves in this article to insulators containing one electron in the lowest excited state, and we shall study the energy spectrum of this single configuration, neglecting perturbations arising from other configurations. As to the method we shall proceed in the following way:

(1) We shall construct orthogonal "atomic" wave functions and express the energy matrix in this vector system.

(2) The energy matrix contains many terms having the periodicity of the lattice and a few which have not; we shall develop a method which takes them both into account.

(3) We shall derive some general results and discuss their consequences.

1. Basic Wave Functions and Energy Matrix

It would no doubt be more satisfactory for insulating crystals, to discuss the Hamiltonian using atomic functions rather than Bloch functions. But this line of attack has been hampered by the fact that atomic functions are not orthogonal. We can, however, build up orthogonal functions having all the advantages of atomic ones by starting out from a Bloch approximation. Let us assume then that a Bloch or Fock method has given us functions $b(\mathbf{k}, \mathbf{x})$ of energy $W(\mathbf{k})$. Then the required functions are

$$a(\mathbf{x}-\mathbf{n}) = 1/(N)^{\frac{1}{2}} \cdot \sum_{\mathbf{k}} \exp\left[-i\mathbf{k},\mathbf{n}\right] b(\mathbf{k},, \mathbf{x}),^{3} \quad (1)$$

where N is the number of cells in the crystal and the $\mathbf{k}$'s are as usual determined by some boundary condition.

Formula (1) applies to any set of Bloch functions, but it might be interesting to get some insight into the structure of the a's. For this purpose let us first make the *ad hoc* assumption (valid for free electrons) that b is of the form

$$b(\mathbf{k},, \mathbf{x}) = \exp\left[i\mathbf{k},\mathbf{x}\right] \cdot b(\mathbf{x}),$$

where the periodic factor $b(\mathbf{x})$ is independent of $\mathbf{k}$. Then we find explicitly:

* I want to express my thanks to Princeton University for the grant of its Swiss-American Exchange Fellowship for the year 1936–37.

[1] J. Frenkel, Phys. Rev. **17**, 17 (1931); Physik. Zeits. Sowjetunion **9**, 158 (1936); Physik. Zeits. Sowjetunion **8**, 185 (1935).

[2] J. C. Slater and W. S. Shockley, Phys. Rev. **50**, 705 (1936).

[3] The unit of length adopted in this article is the elementary translation in the direction of each of the crystal axes. In some deductions the crystal is assumed to be simple cubic, but this could easily be removed.

$$a(\mathbf{x}-\mathbf{n}) = \frac{\sin\pi(x_1-n_1)\sin\pi(x_2-n_2)\sin\pi(x_3-n_3)}{\pi^3(x_1-n_1)(x_2-n_2)(x_3-n_3)} b(\mathbf{x}),$$

which gives us the desired concentration around n_1, n_2, n_3. Their orthogonality is immediately seen because

$$\int_{-\infty}^{+\infty} \frac{\sin\pi x}{x}\frac{\sin\pi(x+n)}{x+n}|b(x)|^2dx = (-)^n/n\cdot\int_{-\infty}^{+\infty}\sin^2\pi x\,|b(x)|^2\left(\frac{1}{x}-\frac{1}{x+n}\right)dx. = 0.$$

These new functions provide us also with a method to orthogonalize standard atomic functions without doing very much harm to them. For this purpose let us take atomic functions $\psi(\mathbf{x}-\mathbf{m})$ and build up from them running waves:

$$b(\mathbf{k},\mathbf{x}) = A(\mathbf{k})/(N)^{\frac{1}{2}}\sum_m \exp[i\mathbf{km}]\cdot\psi(\mathbf{x}-\mathbf{m}). \quad (2)$$

The formula without A gives orthogonal b's, but they are not normalized, if the ψ's are not orthogonal. We find for $1/A^2$:

$$1/A^2(\mathbf{k}) = \sum_n \exp[i\mathbf{kn}]\cdot\int\psi^*(\mathbf{x})\psi(\mathbf{x}-\mathbf{n})d\tau. \quad (3a)$$

From this we can find in principle the value of $A(\mathbf{k})$:

$$A(\mathbf{k}) = \sum_n \exp[-i\mathbf{kn}]\cdot a(\mathbf{n}), \quad (3b)$$

where the a's depend on the integrals in (3a). Putting (3b) and (2) into (1), we find

$$a(\mathbf{x}-\mathbf{m}) = \sum_n a(\mathbf{n}-\mathbf{m})\psi(\mathbf{x}-\mathbf{n}), \quad (4)$$

which is the required result. If, in particular, the integrals are small and fall off slowly, then we can write

$$a(\mathbf{x}-\mathbf{m}) = \psi(\mathbf{x}-\mathbf{m}) - \tfrac{1}{2}\sum_{n\neq m}\psi(\mathbf{x}-\mathbf{n})\int\psi^*(\mathbf{x}-\mathbf{n})\psi(\mathbf{x}-\mathbf{m})d\tau;$$

but if they fall off rapidly, the integral with nearest neighbors determines the coefficient of all ψ's.

As an example let us carry out the orthogonalization (4) explicitly for a set of atomic functions whose integrals are

$$\int\psi^*(\mathbf{x})\psi(\mathbf{x}-\mathbf{n})d\tau = \exp[-\beta_1|n_1| - \beta_2|n_2| - \beta_3|n_3|].$$

Then we get

$$A(\mathbf{k}) = C(\beta_1, k_1)\cdot C(\beta_2, k_2)\cdot C(\beta_3, k_3),$$

where

$$C(\beta, k) = \left(\frac{1-2e^{-\beta}\cos k + e^{-2\beta}}{1-e^{-2\beta}}\right)^{\frac{1}{2}},$$

from which it follows that the expansion coefficients (4) are

$$a(\mathbf{n}) = c(\beta_1, |n_1|)\cdot c(\beta_2, |n_2|)\cdot c(\beta_3, |n_3|),$$

where

$$c(\beta, 0) = 1/(1-e^{-2\beta})^{\frac{1}{2}}F(-\tfrac{1}{2}, -\tfrac{1}{2}, 1, e^{-2\beta}),$$

$$c(\beta, n) = -1/(1-e^{-2\beta})^{\frac{1}{2}}\frac{1}{2n-1}\frac{(2n-1)!!}{(2n)!!} \times e^{-n\beta}F(-\tfrac{1}{2}, n-\tfrac{1}{2}, n+1, e^{-2\beta}).$$

By $m!!$ is meant $m(m-2)(m-4)\cdots$ down to 1 or 2. F is the hypergeometric function

$$F(a, b, c, x) = 1 + \frac{ab}{1c}x + \frac{a(a+1)b(b+1)}{12c(c+1)}x^2 + \cdots.$$

The F's used here are very slowly varying functions. When x increases from 0 to 1, $F(-\frac{1}{2}, -\frac{1}{2}, 1, x)$ increases monotonely from 1 to $4/\pi = 1.273\cdots$ and $F(-\frac{1}{2}, n-\frac{1}{2}, n+1, x)$ decreases monotonely from 1 to

$$[(2n)!!/(2n+1)!!](4/\pi).$$

In most cases one may replace them by 1.

These two examples should show that the new functions defined by (1) go over into atomic ones for large separation and are their natural substitute at smaller distances.

The building up of many electron wave functions from this basis is discussed elsewhere in full detail (e.g. reference 2, appendix). There is a unique wave function for a full band, but for a state having one excited electron and a hole in its full band, three basic types can be constructed:

(a) The atomic wave functions $A(\mathbf{m}, \mathbf{n})$ built up from our "atomic" one-electron functions $a_{\text{ground}}(\mathbf{x}-\mathbf{s})$ and $a_{\text{exc}}(\mathbf{x}-\mathbf{s})$. They contain no electron in the mth cell and an extra excited electron in the nth cell.

(b) The Bloch wave functions $B(\mathbf{k}_p, \mathbf{k}_q)$ built up from one-electron Bloch functions. They contain no electron of wave vector $\mathbf{k}_p$ and an excited electron of wave vector $\mathbf{k}_q$. Between the two there exists the relation:

$$B(\mathbf{k}_p, \mathbf{k}_q) = 1/N \cdot \sum_n \sum_m \exp\left[i\mathbf{k}_q\mathbf{n} - i\mathbf{k}_p\mathbf{m}\right] \cdot A(\mathbf{m}, \mathbf{n}),$$

$$A(\mathbf{m}, \mathbf{n}) = 1/N \cdot \sum_p \sum_q \exp\left[-i\mathbf{k}_q\mathbf{n} + i\mathbf{k}_p\mathbf{m}\right] \cdot B(\mathbf{k}_p, \mathbf{k}_q). \tag{5}$$

(c) The excitation waves $E(\mathbf{K}_\alpha, \mathfrak{z})$ introduced by Frenkel.[1] They describe like Bloch functions the motion of the center of gravity of electron and hole by a wave vector $\mathbf{K}_\alpha$; but as in (a), the two bodies have no relative motion and their distance is given by a space vector $\mathfrak{z}$. Their relation to A and B is

$$E(\mathbf{K}_\alpha, \mathfrak{z}) = 1/(N)^{\frac{1}{2}} \cdot \sum_n \exp\left[i\mathbf{K}_\alpha\mathbf{n}\right] \cdot A(\mathbf{n} - \tfrac{1}{2}\mathfrak{z}, \mathbf{n} + \tfrac{1}{2}\mathfrak{z}), \tag{6}$$

$$B(\mathbf{k}_p, \mathbf{k}_q) = 1/(N)^{\frac{1}{2}} \cdot \sum_{\mathfrak{z}} \exp\left[i\tfrac{1}{2}(\mathbf{k}_p + \mathbf{k}_q)\mathfrak{z}\right] \cdot E(\mathbf{k}_q - \mathbf{k}_p, \mathfrak{z}). \tag{7}$$

The energy matrix is most easily calculated in terms of the A's. We start out from the Hamiltonian

$$H = \sum_i T_i + \sum_{i<k} \frac{e^2}{d \cdot |\mathbf{x}_i - \mathbf{x}_k|} - e \cdot \sum_i \sum_m V(\mathbf{x}_i - \mathbf{m}) \tag{8}$$

$$+ \text{terms independent of } \mathbf{x}_i,$$

where T_i is the kinetic energy of the ith electron, d the length of the elementary crystal-translation in cm, and $\mathbf{x}_i$ the coordinates of the ith electron. The energy is written here in such a form, that only the electrons of the band under consideration are treated as such; the influence of the others is only taken into account through the screened potential $V(\mathbf{x}-\mathbf{m})$ arising from the mth cell.

If one works out the energy matrix $(\mathbf{m}, \mathbf{n}|H|\mathbf{m}', \mathbf{n}')$ one finds a very great number of terms, which depend on $\mathbf{m}, \mathbf{n}, \mathbf{m}', \mathbf{n}'$ only through the combination $\mathbf{m}-\mathbf{m}'$ and $\mathbf{n}-\mathbf{n}'$. The physical reason for this fact is that a transition of the excited electron (or the hole among the unexcited ones) from one cell to another takes place in a field which is very nearly periodic, i.e., the transition $\mathbf{n}\to\mathbf{n}'$ and $\mathbf{n}+\mathbf{s}\to\mathbf{n}'+\mathbf{s}$ must have the same matrix element. In addition, one finds for those periodic terms that if both vectors $\mathbf{m}-\mathbf{m}'$ and $\mathbf{n}-\mathbf{n}'$ are different from zero, the matrix element vanishes; one can therefore distinguish between transitions of the excited electron and transitions of its hole. There remain, however, two integrals which do not follow these two rules; if we write them out explicitly H takes the form:

$$\begin{aligned}(\mathbf{m}, \mathbf{n}|H|\mathbf{m}', \mathbf{n}') &= \delta_{mm'} H_{\text{exc}}(\mathbf{n}-\mathbf{n}') \\ &- \delta_{nn'} H_g(\mathbf{m}-\mathbf{m}') - \int\int a_g^*(\mathbf{x}-\mathbf{m}) a_e^*(\mathbf{x}'-\mathbf{n}) \\ &\times \frac{e^2}{d \cdot |\mathbf{x}-\mathbf{x}'|} a_g(\mathbf{x}-\mathbf{m}') a_e(\mathbf{x}'-\mathbf{n}') d\tau d\tau' \\ &+ \int\int a_g^*(\mathbf{x}-\mathbf{m}) a_e^*(\mathbf{x}'-\mathbf{n}) \frac{e^2}{d \cdot |\mathbf{x}-\mathbf{x}'|} \\ &\times a_g(\mathbf{x}'-\mathbf{m}') a_e(\mathbf{x}-\mathbf{n}') d\tau d\tau'.\end{aligned} \tag{9}$$

These integrals evidently come from the Coulomb attraction between the electron and its hole; they depend on the relative position of the two particles.

There is, however, no term depending on their absolute position in space. We can separate out therefore the motion of their center of gravity and go over to excitation waves.

2. Discussion and Approximate Solution of the Matrix Equation

There are two limiting cases in which (9) is diagonalized immediately. The one is the "infinite separation," where only diagonal terms remain appreciable and either atomic wave functions or excitation waves may be used. The second, which we may call "negligible correlation," occurs when the two integrals can be omitted; then Bloch functions are correct.

This last knowledge can be exploited further, because it means that a Bloch approximation

can take care of all terms H^+ in H, except the two integrals. Let us assume in the future that such a calculation has been carried out and that we know the energy as a function of the wave vector for both bands:

$$(\mathbf{k}_p, \mathbf{k}_q|H^+|\mathbf{k}_{p'}, \mathbf{k}_{q'}) = \delta_{pp'}\delta_{qq'}[W_e(\mathbf{k}_q) - W_g(\mathbf{k}_p)], \quad (10)$$

and therefore, transforming with (5) and comparing with (9):

$$H_e(\mathbf{n}-\mathbf{n}') = 1/N \cdot \sum_q \exp[i\mathbf{k}_q(\mathbf{n}-\mathbf{n}')]W_e(\mathbf{k}_q),$$

$$H_g(\mathbf{m}-\mathbf{m}') = 1/N \cdot \sum_p \exp[-i\mathbf{k}_p(\mathbf{m}-\mathbf{m}')]W_g(\mathbf{k}_p).$$

Equation (9) is now identical with Eq. (21) in Slater's paper,[4] but looks somewhat simpler, because the spin of the electron is assumed to be unchanged in all states involved, and the last term in his expression arising from the change in the self-consistent field is absorbed in W_e and W_g, but a further transformation is in some cases very useful. Let H_e operate on an arbitrary function $F(\mathbf{n})$ satisfying our boundary conditions:

$$F(\mathbf{n}) = \sum_\kappa \exp[i\kappa\mathbf{n}]G(\kappa).$$

We then get

$$\sum_{n'} H_e(\mathbf{n}-\mathbf{n}')F(\mathbf{n}') = 1/N \cdot \sum_q \sum_\kappa \sum_{n'} \exp[i\mathbf{k}_q\mathbf{n} + \mathbf{n}'(\kappa - \mathbf{k}_q)]W_e(\mathbf{k}_q)G(\kappa)$$

$$= \sum_\kappa W_e(\kappa)\exp[i\kappa\mathbf{n}]G(\kappa) = W_e\left(-i\frac{\partial}{\partial n_1}, -i\frac{\partial}{\partial n_2}, -i\frac{\partial}{\partial n_3}\right)\cdot F(\mathbf{n}).$$

This gives for (9)

$$(\mathbf{m}, \mathbf{n}|H|\mathbf{m}', \mathbf{n}') = \delta_{mm'}\delta_{nn'}\left[W_e\left(-i\frac{\partial}{\partial \mathbf{n}}\right) - W_g\left(-i\frac{\partial}{\partial \mathbf{m}}\right)\right] - \int\int a_g^*(\mathbf{x}-\mathbf{m})a_e^*(\mathbf{x}'-\mathbf{n})\frac{e^2}{d\cdot|\mathbf{x}-\mathbf{x}'|}$$

$$\times a_g(\mathbf{x}-\mathbf{m}')a_e(\mathbf{x}'-\mathbf{n}')d\tau d\tau' + \int\int a_g^*(\mathbf{x}-\mathbf{m})a_e^*(\mathbf{x}'-\mathbf{n})\frac{e^2}{d\cdot|\mathbf{x}-\mathbf{x}'|}a_g(\mathbf{x}'-\mathbf{m}')a_e(\mathbf{x}-\mathbf{n}')d\tau d\tau'. \quad (9a)$$

In accordance with the remark at the end of part 1, we now go over to excitation waves as basic wave functions and write our solution in the form

$$\psi = \sum_\beta U_a(\mathfrak{k})E(\mathbf{K}_a, \mathfrak{k}). \quad (11)$$

Then from (6) and (9a), $U(\mathfrak{k})$ has to satisfy the equation

$$\left\{\exp[-i\tfrac{1}{2}\mathbf{K}_a\mathfrak{k}]\cdot W_e\left(-i\frac{\partial}{\partial\mathfrak{k}}\right)\cdot\exp[+i\tfrac{1}{2}\mathbf{K}_a\mathfrak{k}]\right.$$

$$-\exp[+i\tfrac{1}{2}\mathbf{K}_a\mathfrak{k}]\cdot W_g\left(-i\frac{\partial}{\partial\mathfrak{k}}\right)\cdot\exp[-i\tfrac{1}{2}\mathbf{K}_a\mathfrak{k}]$$

$$\left.+\phi(\mathfrak{k})\right\}U(\mathfrak{k}) = E\cdot U(\mathfrak{k}). \quad (12)$$

$\phi(\mathfrak{k})$ is a potential which is very much like $-e^2/(d\cdot|\mathfrak{k}|)$ for large values of $|\mathfrak{k}|$, but has no singularity at $\mathfrak{k}=0$. For simplicity the non-diagonal Coulomb terms have been omitted, for they do not influence the discussion given below and they are so small that they may be neglected in an approximate solution.

The transformation of difference operators into differential operators, as has been done in our transformation from (9) to (9a) and (12), is not much favored by mathematicians. There exists however a book by Davis[5] in which this point of view is adopted, although it is developed for one independent variable only. With this restriction, the theorems on pp. 100–105 of his book ensure that our operator has a sense if $W(\mathbf{k})$ is an entire function of its argument, i.e. defined and regular for complex values of $\mathbf{k}$. Its equivalence with a difference operator may then be shown from the operational identity

[4] J. C. Slater, this issue.

[5] H. D. Davis, *The Theory of Linear Operators* (Bloomington, Ind. 1936).

$$\exp[-i\mathbf{k}\beta]\cdot W\left(-i\frac{\partial}{\partial\beta}\right)\exp[+i\mathbf{k}\beta] = W\left(\mathbf{k}-i\frac{\partial}{\partial\beta}\right), \quad (13)$$

which, because of the periodicity of W, gives

$$W\left(-i\frac{\partial}{\partial\beta}\right)\exp[2\pi i\beta_\nu] - \exp[2\pi i\beta_\nu]W\left(-i\frac{\partial}{\partial\beta}\right)=0, \quad (13a)$$

i.e., a solution U may be multiplied with an arbitrary function having the periodicity of the lattice, and still remains a solution.

Before using (12) for an approximation method, let us first find some general properties of the U's and their energy spectrum.

For large values of β, all the Coulomb integrals and therefore ϕ vanish and the solution must then have exponential character with some complex wave vector $\mathbf{k}=\boldsymbol{\kappa}_1+i\boldsymbol{\kappa}_2$. Remembering (13) we get the corresponding energy from (12):

$$E=W_e(\tfrac{1}{2}\mathbf{K}_a+\boldsymbol{\kappa}_1+i\boldsymbol{\kappa}_2)-W_g(\tfrac{1}{2}\mathbf{K}_a-\boldsymbol{\kappa}_1-i\boldsymbol{\kappa}_2). \quad (14)$$

$\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ are restricted by the condition that E be real. Now we know in advance that E is real, periodic and bounded for $\boldsymbol{\kappa}_2=0$. Therefore, its gradient vanishes at least at two points within one period and we can conclude from general mapping properties that lines of possible complex values of $\mathbf{k}$ start out from these points. The corresponding energies go upwards from the maxima, and downwards from the minima. Fig. 1 is a simple one-dimensional illustration of this situation. W_e-W_g is taken to be $1-\cos k$. The possible complex values of k are then $n\pi+i\kappa$ and their energy is $-\cosh\kappa+1$, if n is even, and $+\cosh\kappa+1$ if n is odd. However, two restrictions reduce the number of these new levels considerably. First of all, the Coulomb field is an attractive field and it can never raise any level; thus there will be no state above the Bloch band. Secondly, a solution of (12) with complex wave vector will in general be unbounded and the condition of boundedness coincides with the condition that $U(\beta)$ vanish exponentially at large distances. This restriction will, for each $\mathbf{K}_a$, select a number of discrete levels out of all possible ones.

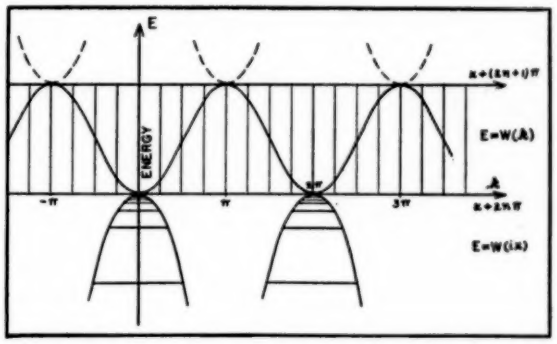


FIG. 1. Energy and wave vector for an excited electron in a crystal.

If we want more detailed information in practical cases we may turn to an atomic picture for the lowest discrete states[6] and to a Bloch picture for the continuum. But for the bulk of the discrete states and the lower continuum there exists a better alternative, namely a local exploration of the Brillouin zone, i.e., if we are interested in a particular region in the Brillouin zone we develop the function $W(\mathbf{k})$ around that point into a power series and take the first two terms or so. The point to be taken in our case is, of course, the bottom of the band where the continuum and the discrete spectrum meet. For practical reasons, let us also assume that the crystal is symmetric enough to allow for a spherically symmetric development of the form

$$\begin{aligned} W_e(\mathbf{k}) &= w_e+A_e(k_1^2+k_2^2+k_3^2),\\ W_g(\mathbf{k}) &= w_g+A_g(k_1^2+k_2^2+k_3^2). \end{aligned} \quad (15)$$

Assuming $A_e-A_g>0$,[7] approximating provisionally $\phi(\beta)$ by $e^2/d\cdot|\beta|$, and making the substitution

$$U(\beta)=\exp\left[\frac{A_e+A_g}{A_e-A_g}i\tfrac{1}{2}\mathbf{K}_a\beta\right]\cdot u(\beta) \quad (16)$$

we find for u the differential equation

$$\Delta u(\beta)+\left(\varepsilon+\frac{2\gamma}{|\beta|}\right)u(\beta)=0 \quad (17)$$

with

$$\gamma=e^2/(2d(A_e-A_g)) \quad (18)$$

[6] R. Peierls, Ann. d. Physik **13**, 905 (1932). A. v. Hippel, Zeits. f. Physik **101**, 680 (1936).

[7] If this condition is reversed, the minimum energy in the Brillouin zone does not lie at (0,0,0), but e. g. at (π,π,π) and one has to study $W(\boldsymbol{\pi}-i\partial/\partial\beta)$. From (13) we see that this gives simply an extra factor -1 for the wave function U, if we pass from one cell to the next.

and $E=w_e-w_g-\mathbf{K}_\alpha^2(A_eA_g)/(A_e-A_g)$

$$+(A_e-A_g)\mathcal{E}. \quad (19)$$

Equation (17) is the Schrödinger equation for the hydrogen atom. It admits for $\mathcal{E}$ all positive eigenvalues and the negative eigenvalues $-\gamma^2/n^2$. One easily verifies that the positive eigenvalues coincide with the Bloch band, while the levels

$$E_n(\mathbf{K}_\alpha)=W_0(\mathbf{K}_\alpha)-\frac{e^4}{4d^2(A_e-A_g)}\frac{1}{n^2} \quad (20)$$

give us the discrete spectrum. The multiplicity of the states belonging to E_n is n^2 as in atomic spectra.

The expressions (19) and (20) are roughly correct as long as approximation (15) holds, i.e., as long as the distance of the level from W_0 is small compared with the width of the whole Bloch band (Fig. 1). Secondly, no negative energy level can lie below the excitation energy of the electron within its own atom; the distance of E from W_0 must therefore also be small compared with the distance of this level. But if one wants to complete these calculations, by a perturbation method, one must remove this second restriction and cut off the Coulomb potential at the origin in a suitable way (Fig. 2). In this case one can at least make the first restriction somewhat less rigid, for one can show that the first order perturbation of the infinite differential operator W exists. It is to be hoped that the perturbation method itself will also converge.

Let us now proceed to establish the chief property of these discrete states, namely that they carry no current.

The current operator in our system is obtained by transforming it from its expression in a Bloch system:

$$(\mathbf{k}_p, \mathbf{k}_q|\,\mathbf{J}\,|\mathbf{k}_{p'}, \mathbf{k}_{q'})$$
$$=-e/\hbar\delta_{pp'}\delta_{qq'}[\nabla W_e(\mathbf{k}_q)-\nabla W_g(\mathbf{k}_p)].$$

The same argument that led from (10) to (9a) and (12) gives now

$$\mathbf{J}U(\mathfrak{z})=-e/\hbar$$
$$\Big\{\exp[-i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{z}]\cdot\nabla W_e\Big(-i\frac{\partial}{\partial\mathfrak{z}}\Big)\cdot\exp[+i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{z}]$$
$$-\exp[+i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{z}]\cdot\nabla W_g\Big(-i\frac{\partial}{\partial\mathfrak{z}}\Big)\cdot\exp[-i\tfrac{1}{2}\mathbf{K}_\alpha\mathfrak{z}]\Big\}$$
$$\times U(\mathfrak{z}). \quad (21)$$

It can easily be shown that apart from surface integrals which vanish if they are extended over the whole crystal, $\int_\Omega U^*(\mathfrak{z})\mathbf{J}U(\mathfrak{z})d\tau_\mathfrak{z}$ is equal to another volume integral which we denote by $\int_\Omega\{U^*\mathbf{J}U\}d\tau_\mathfrak{z}$. One can easily give rules how to build it up from a given $\mathbf{J}$, but they are somewhat lengthy. In this connection, it is sufficient to know that this new expression satisfies a Green's theorem

$$-e\int_V(U^*(\mathfrak{z})WU(\mathfrak{z})-U(\mathfrak{z})WU^*(\mathfrak{z}))d\tau_\mathfrak{z}$$
$$=i\hbar\int_S\{U^*\mathbf{J}U\}d\mathbf{s}_\mathfrak{z},$$

where W stands for the whole differential operator in (12) and the surface integral on the right-hand side is extended over the boundary of the volume on the left. This gives

$$(-e)(d/dt)\int_V U^*(\mathfrak{z})U(\mathfrak{z})d\tau_\mathfrak{z}$$
$$+\int_S\{U^*\mathbf{J}U\}d\mathbf{s}_\mathfrak{z}=0.$$

As in the case of atoms this law is sufficient to prove that if U vanishes exponentially at infinity, then

$$\int_\Omega U^*(\mathfrak{z})\mathbf{J}U(\mathfrak{z})d\tau_\mathfrak{z}=0$$

if the integral is extended over the whole crystal. The true expression is however $\sum_\mathfrak{z}U^*(\mathfrak{z})\mathbf{J}U(\mathfrak{z})$; but one can dispose of the arbitrary periodic factor of U in such a way that

$$\sum_\mathfrak{z}U^*(\mathfrak{z})U(\mathfrak{z})=\int_\Omega U^*(\mathfrak{z})U(\mathfrak{z})d\tau_\mathfrak{z}$$

and $|\sum_\mathfrak{z}U^*(\mathfrak{z})\mathbf{J}U(\mathfrak{z})|\leqq|\int_\Omega U^*(\mathfrak{z})\mathbf{J}U(\mathfrak{z})d\tau_\mathfrak{z}|+\epsilon,$

which proves the theorem.

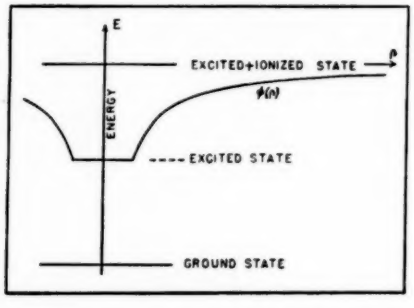


FIG. 2. Potential energy of electron and hole as a function of their distance.

3. General Survey of the Results

The results of the above calculations suggest a rather fundamental change in the picture used to describe the behavior of electrons in crystals.

Since there has been so much talk about discrete levels, one should perhaps first remember that there are no discrete absorption lines to be expected in any crystalline spectrum, because, as in molecules, there is a band structure connected with each electronic energy level. The value of this term in the approximation (19) is

$$-\frac{A_e A_g \mathbf{K}_\alpha^2}{(A_e - A_g)}.$$

The analogy of $\mathbf{K}_\alpha$ to the vibrational quantum number of a molecular band spectrum becomes evident, if we remember Frenkel's selection rule.[1] It states that if a light wave of wave vector κ excites a state belonging to the excitation waves of wave vector $\mathbf{K}_\alpha$ the difference $\kappa - \mathbf{K}_\alpha$ must appear as wave vector of a simultaneously absorbed or emitted lattice wave. Since for visible light κ is very small, $\mathbf{K}_\alpha$ may be considered as the wave number of this lattice vibration.

The truly electronic spectrum underlying the band spectrum is obtained by putting $\mathbf{K}_\alpha = 0$ in the energy expression (20). It is divided into two parts:

First there comes a lower part consisting of discrete states. The electron does not escape its hole and no photocurrent can be observed. The very lowest states are widely spaced and correspond to excitation of an electron within its cell or to some direct neighbor.[6] As one goes higher up this individual character gets lost; the spacing gets narrower and the electron moves in an orbit determined by the Coulomb field of its hole.[8] For a finite crystal with N cells, the number of such states is of the order $N^{3/5}$.

Then there follows a continuum which is just the Bloch band. Electron and hole move independently and a current may be observed. But in these states the electron cannot well approach its hole and optical transitions into them are therefore unlikely. For our finite crystal, the number of such states is of the order N.

As far as can be told, experiment is in agreement with these views.[9]

In the practical application of these results one should keep in mind some serious restrictions which underlie the calculations.

This paper deals with the excited states of a single electron only. It will still be good for the excitation of a number of electrons small compared with the number of cells, but after this qualitative changes may occur; the same is true for conducting crystals.

Secondly these calculations deal with virtual excited states of an unexcited crystal rather than with truly excited states. And the virtual states considered are those compatible with the Frank-Condon principle. Each electronic excitation will however be followed by rearrangements of the lattice and energy dissipation. Many authors believe that this will result in an actual "trapping" of the electron in its new position.[1, 5]

One should think that these facts have no influence on the absorption spectrum of a crystal, but it has been pointed out to me by Morse that, even for this case, the list of levels given above may be incomplete, because there is a possibility of weak absorptions violating the Frank-Condon principle. This is especially true for regions where no other absorption is possible. It is perhaps the explanation for the photoelectric absorption of many crystals on the long wave-length foot of their ultraviolet band.

Finally I want to express my thanks to Professor E. Wigner, who has suggested to me the present problem, and to Professor J. C. Slater, Dr. C. Herring, Dr. F. Smithies and many others for valuable advice. I also want to thank Mr. R. C. Herman for his help in completing this paper.

[8] This distinction is well known in the theory of electrolytes. Our lowest levels correspond to the classical theory of weak electrolytes, the higher ones to the theory of Debye-Hückel. This analogy may perhaps be extended if one considers the motion of "trapped" electrons.

[9] R. Hilsch and R. W. Pohl, Zeits. f. Physik **59**, 812 (1930); E. G. Schneider and H. M. O'Bryan, Phys. Rev. **51**, 293 (1937); B. Gudden and R. Pohl, Zeits. f. Physik **17**, 331 (1923); F. C. Nix, Rev. Mod. Phys. **4**, 723 (1932); A. L. Hughes, Rev. Mod. Phys. **8**, 294 (1936).

The Theory of Ferromagnetism: Lowest Energy Levels

J. C. SLATER
*Institute for Advanced Study, Princeton, New Jersey**

(Received May 7, 1937)

The theory of ferromagnetism is formulated both by the method of energy bands and the method of spin waves, and the connection between them is investigated. The particular problem considered is that where N–1 electrons in a band have one spin, one electron the opposite spin. Each method yields a set of unperturbed wave functions, each of which can be expressed in terms of the other. Neither represents the true situation, which must be found by perturbation theory. The final result is that below the continuum of levels indicated by the energy band theory there is a set of discrete levels. The lowest of these is essentially a spin wave, and is the state ordinarily occupied. For investigating the temperature variation of magnetization, the spin wave theory should be used. On the other hand, the energy band theory leads to approximately correct conclusions as to which elements should be ferromagnetic. The formulation is in terms of orthogonal atomic functions, which have advantages compared with the nonorthogonal functions usually employed. In the limit of broad energy bands, and nonferromagnetic substances, the discrete levels approach the continuum, and their wave functions represent a situation in which the electron of negative spin can wander to a considerable distance from the positive ion which it has left, but not to infinite distance.

THE theory of ferromagnetism has been treated by several distinct methods. In the first place, Heisenberg's[1] original theory was based on the Heitler and London method in molecular structure, each electron being localized on a particular atom, and Bloch's[2] theory of spin waves, and Bethe's[3] extension of that theory, were based on Heisenberg's work. On the other hand, Bloch[4] worked out a theory of the ferromagnetism of free electrons, based on the theory of energy bands. This type of theory has been adapted by the present writer[5] to a discussion of the ferromagnetism of nickel. Both methods have advantages and disadvantages. Heisenberg's type of theory has in the first place the great advantage that it makes direct connection with the familiar Weiss theory of ferromagnetism, though it does not agree with that theory in all details. Furthermore, it reduces to the proper limiting values when the atoms of the crystal are removed to infinite separation from each other. It has, however, three great disadvantages. The first is that it does not change over continuously to a theory of metallic conduction electrons, when proper changes are made in the various parameters involved. It is known that this is a result of the failure to include polar states in the formulation, states where atoms can carry a net charge. The second disadvantage is that it does not lead in any very clear way to an understanding of why some elements are ferromagnetic, others not. The exchange integral which determines ferromagnetism must be of opposite sign to its normal value for ferromagnetism to occur, and while plausible reasons have been given to suppose that it actually has this sign in the ferromagnetic elements, these reasons have not, as far as the writer knows, been strengthened by an actual calculation of the exchange integral. The third disadvantage is a rather technical, but none the less a very important, one: the wave functions with which one operates are not orthogonal, with the result that there are many exchange integrals between three and more electrons, which are generally neglected in the simple form of the theory, but which are actually not negligible. Van Vleck[6] has recently shown how to handle this lack of orthogonality, but even if it can be taken care of, it seriously complicates the calculations.

The theory based on energy bands likewise has its advantages and disadvantages. Greatest of the advantages is probably the fact that it makes direct connection with the theory of conductivity, of cohesion, and of other metallic properties, all now handled most satisfactorily

* On leave from the Massachusetts Institute of Technology, Cambridge, Mass.

[1] W. Heisenberg, Zeits. f. Physik **49**, 619 (1928).
[2] F. Bloch, Zeits. f. Physik **61**, 206 (1930).
[3] H. Bethe, Zeits. f. Physik **71**, 205 (1931).
[4] F. Bloch, Zeits. f. Physik **57**, 545 (1929).
[5] J. C. Slater, Phys. Rev. **49**, 537, 931 (1936).
[6] J. H. Van Vleck, Phys. Rev. **49**, 232 (1936).

by energy bands. Secondly, it leads to a clear understanding, based on approximate calculations, of the reason why ferromagnetism occurs where it does in the periodic table. Thirdly, it is based throughout on orthogonal functions, greatly simplifying the calculations. Finally, the theory of energy bands has reached a stage where fairly satisfactory numerical calculations can be made, so that it is not hard to make our information fairly quantitative. On the other hand, this theory also has serious disadvantages. It leads to quite wrong values in the limiting case of large separation of the atoms, and this is likely to be particularly serious, since the energy bands responsible for ferromagnetism are narrow, corresponding to atomic orbits which are widely separated in proportion to their size. Furthermore, it makes no connection with Weiss's theory, and with the valuable way of visualizing the problem which that theory provides. We must conclude that neither method by itself is very satisfactory.

This is no newly appreciated situation. The writer in 1930[7] called attention to it, and pointed out that energy band functions on the one hand, or atomic functions of Heisenberg's type, augmented by polar states, on the other, formed two possible starting points for a perturbation calculation. Either set of functions can be written as linear combinations of the functions of the other set. A correct linear combination of the functions of either set will then yield a more satisfactory approximation than functions of either set separately. In other words, we can make a theory more satisfactory than either Heisenberg's or the energy band theory, and the final result will be independent of which we use as a starting point. These remarks were illustrated at the time by the hydrogen molecule H_2, which can be treated by analogous theories, but it was not possible then to make detailed calculations for the problem of ferromagnetism, which was not the primary concern of the paper mentioned. Since then, Schubin and Wonsowsky[8] have extended the use of polar states, obtaining results analogous to some of those in the present paper. The writer and Shockley[9] recently discussed the very similar problem of absorption of light by the alkali halides, and the method there used can be adapted to ferromagnetism. Furthermore, recent work of Wannier[10] has provided us with orthogonal atomic functions, which simplify the problem greatly. It therefore seems worth while to approach the general theory again, with a view to finding the best solution of the problem intermediate between the Heisenberg and the energy band type of theory. The present paper contains a first step in this direction. We consider only a single energy band, which in its lowest state contains just one electron per atom of one spin, none of the other spin. And we consider just the case where one electron only is disturbed from this state, reversing its spin, raising the energy of the system if it is ferromagnetic, or lowering it if it is not. This is essentially the same problem considered in Bloch's theory of spin waves,[2] but supplemented by polar states, allowing the connection with the theory of energy bands to be made. Though this case is much simpler than the real situation, still it is instructive. The result is a theory which reduces precisely in one limit to the theory of energy bands, and in another limit to something very much like the Heisenberg theory, possessing all the advantages of that theory, but the additional great advantage of dealing only with orthogonal functions. The intermediate case, representing the actual state of affairs, proves to resemble the theory of energy bands in the criterion for ferromagnetism, verifying in a general way the results of the author's calculation[5] that the elements of the iron group should be ferromagnetic. But the actual lowest states of the system are described by something much more like Bloch's spin waves, so that Bloch's calculation of the temperature variation of saturation magnetization at low temperatures should be correct. The present paper, like Bloch's, is limited to low temperatures, since we can only treat the case where one electron has its spin reversed, or practically the case where there are so few reversed spins that they rarely interact with each other.

[7] J. C. Slater, Phys. Rev. **35**, 509 (1930).
[8] S. Schubin and S. Wonsowsky, Proc. Roy. Soc. **A145**, 159 (1934).
[9] J. C. Slater and W. Shockley, Phys. Rev. **50**, 705 (1936).
[10] G. Wannier, Phys. Rev. (preceding paper, this issue).

The Unperturbed Wave Functions

We shall start with the method of energy bands. Imagine a periodic potential $U(\mathbf{r})$, where $\mathbf{r}$ is a vector coordinate. We shall later determine $U(\mathbf{r})$ by a Fock-Hartree self-consistent field method. We shall now imagine that we have the exact solutions for the motion of an electron in this potential. We consider only one energy band, containing N energy levels of each spin, where N is the number of atoms in the crystal (or in the repeating region of space, using periodic boundary conditions). Let the wave function associated with the wave vector $\mathbf{k}$ be $b(\mathbf{k}, \mathbf{r})$, so that it is multiplied by $\exp(i\mathbf{k}\cdot\mathbf{R})$ in traveling from one cell to another distant by the vector $\mathbf{R}$. Let the energy as a function of $\mathbf{k}$ be $W(\mathbf{k})$. Since we are dealing with only one band, we may always imagine $\mathbf{k}$ to be in the central cell of $\mathbf{k}$ space. Using atomic units, the functions then satisfy the Schrödinger equation

$$(-\nabla^2+U(\mathbf{r}))b(\mathbf{k},\mathbf{r})=W(\mathbf{k})b(\mathbf{k},\mathbf{r}). \tag{1}$$

We shall often have occasion to use atomic functions, rather than the periodic functions $b(\mathbf{k}, \mathbf{r})$. For this purpose we shall use Wannier's orthogonal functions:[10]

$$a(\mathbf{r}-\mathbf{R})=1/(N)^{\frac{1}{2}}\sum_{\mathbf{k}}\exp\left[(-i\mathbf{k}\cdot\mathbf{R})\right]b(\mathbf{k},\mathbf{r}). \tag{2}$$

Here $\mathbf{R}$ is the radius vector to one of the nuclei, and the summation is over the N vectors $\mathbf{k}$ in the central cell. We can easily prove that the a's are orthogonal, and normalized if the b's are normalized:

$$\int a^*(\mathbf{r}-\mathbf{R})a(\mathbf{r}-\mathbf{R}')d\tau=(1/N)\sum_{\mathbf{k},\,\mathbf{k}'}\exp\left[i(\mathbf{k}\cdot\mathbf{R}-\mathbf{k}'\cdot\mathbf{R}')\right]\int b^*(\mathbf{k},\mathbf{r})b(\mathbf{k}',\mathbf{r})d\tau$$
$$=(1/N)\sum_{\mathbf{k}}\exp\left[i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')\right], \tag{3}$$

on account of the orthogonality and normalization of the b's. But it is a property of the reciprocal lattice that the expression on the right side of (3) is zero if $\mathbf{R}-\mathbf{R}'$ is a vector between two lattice points which is different from zero, but is unity if $\mathbf{R}-\mathbf{R}'$ is zero. Hence the orthogonality and normalization of the a's are proved. In the future we shall often use the property of summing over $\mathbf{k}$ which we have just mentioned, and also the other property that the same summation mentioned in (3), but now summed over all atoms $\mathbf{R}'$ rather than over all wave vectors $\mathbf{k}$, is zero unless $\mathbf{k}$ equals zero, in which case it is unity.

The summation in (2), by which the a's are expressed as functions of the b's, is such that $a(\mathbf{r}-\mathbf{R})$ is largely concentrated about the nucleus at $\mathbf{R}$, and resembles on that nucleus the atomic function of the isolated atom. The function changes sign on the nearest neighbors, however, so as to result in orthogonality, and in fact oscillates in sign, with rapidly decreasing amplitude, as we go away from the nucleus at $\mathbf{R}$. As the atoms become widely separated, if $U(\mathbf{r})$ approaches the potential of an isolated atom about each nucleus, $a(\mathbf{r}-\mathbf{R})$ approaches the corresponding atomic wave function of the isolated atom. The function $a(\mathbf{r}-\mathbf{R})$ is not conveniently written as the solution of an ordinary Schrödinger equation; we shall instead express its differential equation through the transformation (2), in terms of the Eq. (1) for the functions $b(\mathbf{k}, \mathbf{r})$.

We shall now set up unperturbed wave functions for the whole system, in terms of the functions $b(\mathbf{k}, \mathbf{r})$. First we consider the state where all N electrons have + spins. Then the wave function is simply the determinant of functions b, using all values of $\mathbf{k}$ and the $\mathbf{r}$'s of all electrons:

$$B(0)=1/(N!)^{\frac{1}{2}}\,|\,b(\mathbf{k},\mathbf{r}_i)\,|\,. \tag{4}$$

This state $B(0)$ will be the lowest state in the ferromagnetic case. Next we consider the states where $(N-1)$ electrons have + spin, one has − spin. The process of setting up these states is almost exactly as in the paper on absorption by the alkali halides.[9] We consider in particular a state in which the electron with + spin, and wave vector $\mathbf{k}-\mathbf{K}$, is missing, but in which there

is an electron with $-$ spin, with wave vector $\mathbf{k}$. The corresponding wave function is defined as $B(\mathbf{k}-\mathbf{K}, \mathbf{k})$, and it differs from (4) in that the elements of the row equal to $b_+(\mathbf{k}-\mathbf{K}, \mathbf{r}_i)$ in (4) are missing, and are replaced by the elements $b_-(\mathbf{k}, \mathbf{r}_i)$. There are N^2 such states, corresponding to all N possible values of $\mathbf{k}$ and $\mathbf{K}$. It is easily shown, however, that the matrix components of energy, and of all periodic operators, are zero between states of different $\mathbf{K}$ value. This quantity $\mathbf{K}$ is at once seen to measure the total wave vector of all electrons in the excited state. Thus states with different $\mathbf{K}$ do not combine. Our perturbation problem will be one between the N states of a given $\mathbf{K}$. We shall have to find matrix components of the energy between them, and find the suitable linear combinations to give the correct solutions of the problem.

In addition to these functions B, we shall use excitation functions $F(\mathbf{K}, \mathbf{R})$, similar to the functions E and F introduced in reference (9). We define

$$F(\mathbf{K}, \mathbf{R}) = 1/(N)^{\frac{1}{2}} \sum_{\mathbf{k}} \exp\left[i\mathbf{k}\cdot\mathbf{R}\right] B(\mathbf{k}-\mathbf{K}, \mathbf{k}). \quad (5)$$

By methods similar to those of reference 9, we can show that $F(\mathbf{K}, \mathbf{R})$ can be made up in the following way. Starting with the lowest wave function $B(0)$, we describe the system by atomic functions $a(\mathbf{r}-\mathbf{R}_i)$. Then we remove the electron of $+$ spin from the atom $\mathbf{R}_i$, reverse its spin, and place it on the atom at $\mathbf{R}_i+\mathbf{R}$, producing a positive ion if $\mathbf{R}\neq 0$, though there are no ions produced if $\mathbf{R}=0$. We now combine all such functions from different atoms $\mathbf{R}_i$, multiplying each by a factor $1/(N)^{\frac{1}{2}} \exp i\mathbf{K}\cdot\mathbf{R}_i$, and adding. The function $F(\mathbf{K}, 0)$ then corresponds to Bloch's spin wave, except that our atomic functions a do not agree exactly with the atomic functions of Heisenberg's theory, and the functions $F(\mathbf{K}, \mathbf{R})$ with $\mathbf{R}\neq 0$ are the polar states which must be included to supplement Heisenberg's theory.

Matrix Components of Energy

We must next set up the matrix components of the energy of the whole system. This is most easily done first in the system of B functions, and later we can find the components in the F system by transformation theory. The first step is to define the potential field U. We assume each atom to consist of a nucleus, and enough electrons in closed shells to leave it a singly charged positive ion, so that the one remaining electron which we are considering in our perturbation problem will make it neutral. Let $V(\mathbf{r})$, where $\mathbf{r}$ is the vector position, represent the potential due to the nuclei and closed shells of all the atoms, averaged in a suitable manner. This will be a sum of terms, one from each atom, and each falling off inversely as the distance at a distance from the nucleus, but behaving differently within the closed shells. In order to get $U(\mathbf{r})$ from $V(\mathbf{r})$, we must add the potential of the electrons in the states we are considering. In the first place, if all functions $b(\mathbf{k}, \mathbf{r})$ were occupied with one electron each, the electronic charge density would be $\sum_{\mathbf{k}} b^*(\mathbf{k}, \mathbf{r}) b(\mathbf{k}, \mathbf{r})$, so that we should have to add the potential

$$\text{electron potential} = \int \sum_{\mathbf{k}} b^*(\mathbf{k}, \mathbf{r}_j) b(\mathbf{k}, \mathbf{r}_j) (2/|\mathbf{r}_{ij}|) d\tau_j. \quad (6)$$

This is the potential, at point $\mathbf{r}_i$, of the distributed charge in elements $d\tau_j$, at distance $\mathbf{r}_{ij}$ from $\mathbf{r}_i$. By means of Eq. (2) or its inverse, we can express the electron potential in terms of atomic functions. It is then

$$\text{electron potential} = \int \sum_{s} a^*(\mathbf{r}_j - \mathbf{R}_s) a(\mathbf{r}_j - \mathbf{R}_s) (2/|\mathbf{r}_{ij}|) d\tau_j. \quad (7)$$

The potential (6) or (7) is not exactly the correct one to use, however, for it is the potential of N electrons acting on the electron we are interested in, whereas actually only $N-1$ electrons are left to act on it, and on account of exchange, or the Pauli principle, the one electron missing from the distribution is removed from the immediate neighborhood of $\mathbf{r}_i$, so that it has an important effect on the potential. Remembering the orthogonality of the b's and a's, and the fact that in the lowest state all electrons have the same spin, this exchange charge density, at the point $\mathbf{r}_j$, when the electron we are interested in as at $\mathbf{r}_i$, is given by[11]

[11] J. C. Slater, Rev. Mod. Phys. **6**, 267 (1934).

$$\{\sum_{\mathbf{k}}\sum_{\mathbf{k}'} b^*(\mathbf{k}, \mathbf{r}_i)b(\mathbf{k}', \mathbf{r}_i)b^*(\mathbf{k}', \mathbf{r}_j)b(\mathbf{k}, \mathbf{r}_j)\}/\{\sum_{\mathbf{k}''} b^*(\mathbf{k}'', \mathbf{r}_i)b(\mathbf{k}'', \mathbf{r}_i)\}$$

$$= \{\sum_{s}\sum_{t} a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)\}/\{\sum_{v} a^*(\mathbf{r}_i-\mathbf{R}_v)a(\mathbf{r}_i-\mathbf{R}_v)\}. \quad (8)$$

From the second of expressions (8) we can see particularly easily that the exchange charge density is located in the neighborhood of $\mathbf{r}_i$. If $\mathbf{r}_i$ is in the neighborhood of the nucleus at $\mathbf{R}_i$, then terms like $a(\mathbf{r}_i-\mathbf{R}_s)$ will be much larger for $s=i$ than for any other value of s. Hence in the summations over s, t, and v the leading term will be that for which $s=t=v=i$. Then expression (8) reduces simply to $a^*(\mathbf{r}_j-\mathbf{R}_i)a(\mathbf{r}_j-\mathbf{R}_i)$, the distribution about the nucleus $\mathbf{R}_i$. This approximation becomes exact as the atoms become separated from each other. Now from the exchange charge density (8) we can set up the exchange potential:

exchange potential =

$$\left\{\sum_{s}\sum_{t} a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)\int a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)(2/|\mathbf{r}_{ij}|)d\tau_j\right\}\Big/\left\{\sum_{v} a^*(\mathbf{r}_i-\mathbf{R}_v)a(\mathbf{r}_i-\mathbf{R}_v)\right\}. \quad (9)$$

In terms of these quantities, we then have

$$U(\mathbf{r}_i) = V(\mathbf{r}_i) + \text{electron potential} - \text{exchange potential}. \quad (10)$$

From what we have just seen, this potential U behaves as it should when the atoms become widely separated; that is, it reduces near each nucleus to the potential of that nucleus and the electrons surrounding it, unaffected by the other atoms. Since the electrons all have parallel spins, no correction for correlation is necessary in addition to the exchange correction. Although the potential we have set up is strictly correct only for the lowest state, we use it as well for the states we are particularly interested in, where all electrons but one are parallel, the remaining one antiparallel. This approximation will make a negligible error, and is entirely justified in the present case, though it might not be if we were considering states with many antiparallel spins.

Next we must find the matrix components of energy. The energy operator for the whole system is

$$H = -\sum_i \nabla_i^2 + \sum_i V(\mathbf{r}_i) + \sum_{\text{pairs } ij} 2/|\mathbf{r}_{ij}| + \sum_{\text{pairs } st} 2/|\mathbf{R}_{st}|. \quad (11)$$

The first term is the kinetic energy, the second the potential energy of electrons in the field of the nuclei and inner shells, the third the repulsion of electrons, the fourth the repulsion of the positive ions. Aside from the last term, which does not depend on the electron coordinates at all, we have two types of term: the first and second, a sum of one-electron operators, and the third, a sum of two-electron operators. Remembering that the one-electron functions are orthogonal, the diagonal component of the energy matrix equals the sum of the one-electron matrices between all one-electron functions, plus the sum of two-electron matrices between all pairs of functions, minus the sum of two-electron exchange integrals between all pairs with parallel spins. For the lowest state, then, we have in a simple notation

$$(B(0)|H|B(0)) = \sum_{\mathbf{k}} (b(\mathbf{k}, \mathbf{r})|-\nabla^2+V|b(\mathbf{k}, \mathbf{r}))$$

$$+ \sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_s)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_t)2/|\mathbf{r}_{ij}|d\tau_{ij}$$

$$- \sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)2/|\mathbf{r}_{ij}|d\tau_{ij} + \sum_{\text{pairs } s,\ t} 2/|\mathbf{R}_{st}|, \quad (12)$$

where we have used the functions b for the one-electron operators, the a's for the two-electron ones, for reasons of convenience. We find the Laplacian operators from Eq. (1), substitute for $U(\mathbf{r})$ from Eq. (10), remember that a summation over all values of s and t independently equals twice the summation over all pairs, and finally change all terms over to dependence on the a's rather than the b's. Then we find

$$(B(0)|H|B(0))=E_0=\sum_{\mathbf{k}}W(\mathbf{k})+\sum_{\text{pairs } s,\ t}2/|\mathbf{R}_{st}|$$

$$-\sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_s)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_t)2/|\mathbf{r}_{ij}|d\tau_{ij}$$

$$+\sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)2/|\mathbf{r}_{ij}|d\tau_{ij}. \quad (13)$$

We shall use this energy E_0 simply to establish a zero of energy. If, however, we were interested in cohesion, its variation with lattice spacing would give us the total energy of the crystal. Since there is only one valence electron per atom, the terms other than the first will approximately cancel,[12] leaving the sum of $W(\mathbf{k})$'s as the approximate energy. (Of course, our model is really very deficient for a ferromagnetic substance. The electrons we are considering correspond to the $3d$ electrons in the ferromagnetic elements, and we neglect entirely the valence electrons which produce the binding.) In the limit of large internuclear distances, this relation becomes exact, and since the $W(\mathbf{k})$'s all coincide, and equal the atomic energy, at large distance, this leads to the correct energy for the lowest state of the system.

Next we find the diagonal energy $(B(\mathbf{k}-\mathbf{K}, \mathbf{k})|H|B(\mathbf{k}-\mathbf{K}, \mathbf{k}))$ for one of our states with a reversed spin. Most of the terms will be the same as those already considered in (12) or (13), and we shall simply consider the terms which are different. As far as the one-electron operators are concerned, we must eliminate the terms coming from $b(\mathbf{k}-\mathbf{K}, \mathbf{r})$, and add extra terms coming from $b(\mathbf{k}, \mathbf{r})$. It is more convenient in these expressions to retain the functions b, rather than writing in terms of atomic functions, since $\mathbf{k}$ appears explicitly in the final result. Then we find without complications that the one-electron operators give terms differing from those of (12) by

$$W(\mathbf{k})-W(\mathbf{k}-\mathbf{K})+\sum_{\mathbf{k}'}\int(b^*(\mathbf{k}-\mathbf{K}, \mathbf{r}_1)b(\mathbf{k}-\mathbf{K}, \mathbf{r}_1)-b^*(\mathbf{k}, \mathbf{r}_1)b(\mathbf{k}, \mathbf{r}_1))b^*(\mathbf{k}', r_2)b(\mathbf{k}', \mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$-\sum_{\mathbf{k}'}\sum_{\mathbf{k}''}\int\{(b^*(\mathbf{k}-\mathbf{K}, \mathbf{r}_1)b(\mathbf{k}-\mathbf{K}, \mathbf{r}_1)-b^*(\mathbf{k}, \mathbf{r}_1)b(\mathbf{k}, \mathbf{r}_1))b^*(\mathbf{k}', \mathbf{r}_1)b(\mathbf{k}'', \mathbf{r}_1)$$

$$\times b^*(\mathbf{k}'', \mathbf{r}_2)b(\mathbf{k}', \mathbf{r}_2)\}/\{\sum_{\mathbf{k}'''}b^*(\mathbf{k}''', \mathbf{r}_1)b(\mathbf{k}''', \mathbf{r}_1)\}2/|\mathbf{r}_{12}|d\tau_{12}. \quad (14)$$

The two-electron operators give terms differing by

$$-\sum_{\mathbf{k}'}\int(b^*(\mathbf{k}-\mathbf{K}, \mathbf{r}_1)b(\mathbf{k}-\mathbf{K}, \mathbf{r}_1)-b^*(\mathbf{k}, \mathbf{r}_1)b(\mathbf{k}, \mathbf{r}_1))b^*(\mathbf{k}', \mathbf{r}_2)b(\mathbf{k}', \mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$-\int(b^*(\mathbf{k}, \mathbf{r}_1)b(\mathbf{k}, \mathbf{r}_1)b^*(\mathbf{k}-\mathbf{K}, \mathbf{r}_2)b(\mathbf{k}-\mathbf{K}, \mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$+\sum_{\mathbf{k}'}\int b^*(\mathbf{k}-\mathbf{K}, \mathbf{r}_1)b(\mathbf{k}', \mathbf{r}_1)b^*(\mathbf{k}', \mathbf{r}_2)b(\mathbf{k}-\mathbf{K}, \mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}. \quad (15)$$

[12] J. C. Slater, Rev. Mod. Phys. **6**, 270 (1934).

$$\{\sum_{\mathbf{k}}\sum_{\mathbf{k}'} b^*(\mathbf{k}, \mathbf{r}_i)b(\mathbf{k}', \mathbf{r}_i)b^*(\mathbf{k}', \mathbf{r}_j)b(\mathbf{k}, \mathbf{r}_j)\}/\{\sum_{\mathbf{k}''} b^*(\mathbf{k}'', \mathbf{r}_i)b(\mathbf{k}'', \mathbf{r}_i)\}$$

$$= \{\sum_{s}\sum_{t} a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)\}/\{\sum_{v} a^*(\mathbf{r}_i-\mathbf{R}_v)a(\mathbf{r}_i-\mathbf{R}_v)\}. \quad (8)$$

From the second of expressions (8) we can see particularly easily that the exchange charge density is located in the neighborhood of $\mathbf{r}_i$. If $\mathbf{r}_i$ is in the neighborhood of the nucleus at $\mathbf{R}_i$, then terms like $a(\mathbf{r}_i-\mathbf{R}_s)$ will be much larger for $s=i$ than for any other value of s. Hence in the summations over s, t, and v the leading term will be that for which $s=t=v=i$. Then expression (8) reduces simply to $a^*(\mathbf{r}_j-\mathbf{R}_i)a(\mathbf{r}_j-\mathbf{R}_i)$, the distribution about the nucleus $\mathbf{R}_i$. This approximation becomes exact as the atoms become separated from each other. Now from the exchange charge density (8) we can set up the exchange potential:

exchange potential =

$$\left\{\sum_{s}\sum_{t} a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)\int a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)(2/|\mathbf{r}_{ij}|)d\tau_j\right\}\Big/\left\{\sum_{v} a^*(\mathbf{r}_i-\mathbf{R}_v)a(\mathbf{r}_i-\mathbf{R}_v)\right\}. \quad (9)$$

In terms of these quantities, we then have

$$U(\mathbf{r}_i) = V(\mathbf{r}_i) + \text{electron potential} - \text{exchange potential}. \quad (10)$$

From what we have just seen, this potential U behaves as it should when the atoms become widely separated; that is, it reduces near each nucleus to the potential of that nucleus and the electrons surrounding it, unaffected by the other atoms. Since the electrons all have parallel spins, no correction for correlation is necessary in addition to the exchange correction. Although the potential we have set up is strictly correct only for the lowest state, we use it as well for the states we are particularly interested in, where all electrons but one are parallel, the remaining one antiparallel. This approximation will make a negligible error, and is entirely justified in the present case, though it might not be if we were considering states with many antiparallel spins.

Next we must find the matrix components of energy. The energy operator for the whole system is

$$H = -\sum_i \nabla_i^2 + \sum_i V(\mathbf{r}_i) + \sum_{\text{pairs } ij} 2/|\mathbf{r}_{ij}| + \sum_{\text{pairs } st} 2/|\mathbf{R}_{st}|. \quad (11)$$

The first term is the kinetic energy, the second the potential energy of electrons in the field of the nuclei and inner shells, the third the repulsion of electrons, the fourth the repulsion of the positive ions. Aside from the last term, which does not depend on the electron coordinates at all, we have two types of term: the first and second, a sum of one-electron operators, and the third, a sum of two-electron operators. Remembering that the one-electron functions are orthogonal, the diagonal component of the energy matrix equals the sum of the one-electron matrices between all one-electron functions, plus the sum of two-electron matrices between all pairs of functions, minus the sum of two-electron exchange integrals between all pairs with parallel spins. For the lowest state, then, we have in a simple notation

$$(B(0)|H|B(0)) = \sum_{\mathbf{k}}(b(\mathbf{k}, \mathbf{r})|-\nabla^2+V|b(\mathbf{k}, \mathbf{r}))$$

$$+ \sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_s)a^*(r_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_t)2/|\mathbf{r}_{ij}|d\tau_{ij}$$

$$- \sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)2/|\mathbf{r}_{ij}|d\tau_{ij} + \sum_{\text{pairs } s,\ t} 2/|\mathbf{R}_{st}|, \quad (12)$$

where we have used the functions b for the one-electron operators, the a's for the two-electron ones, for reasons of convenience. We find the Laplacian operators from Eq. (1), substitute for $U(\mathbf{r})$ from Eq. (10), remember that a summation over all values of s and t independently equals twice the summation over all pairs, and finally change all terms over to dependence on the a's rather than the b's. Then we find

$$(B(0)|H|B(0))=E_0=\sum_{\mathbf{k}}W(\mathbf{k})+\sum_{\text{pairs } s,\ t}2/|\mathbf{R}_{st}|$$

$$-\sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_s)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_t)2/|\mathbf{r}_{ij}|d\tau_{ij}$$

$$+\sum_{\text{pairs } s,\ t}\int\int a^*(\mathbf{r}_i-\mathbf{R}_s)a(\mathbf{r}_i-\mathbf{R}_t)a^*(\mathbf{r}_j-\mathbf{R}_t)a(\mathbf{r}_j-\mathbf{R}_s)2/|\mathbf{r}_{ij}|d\tau_{ij}. \quad (13)$$

We shall use this energy E_0 simply to establish a zero of energy. If, however, we were interested in cohesion, its variation with lattice spacing would give us the total energy of the crystal. Since there is only one valence electron per atom, the terms other than the first will approximately cancel,[12] leaving the sum of $W(\mathbf{k})$'s as the approximate energy. (Of course, our model is really very deficient for a ferromagnetic substance. The electrons we are considering correspond to the *3d* electrons in the ferromagnetic elements, and we neglect entirely the valence electrons which produce the binding.) In the limit of large internuclear distances, this relation becomes exact, and since the $W(\mathbf{k})$'s all coincide, and equal the atomic energy, at large distance, this leads to the correct energy for the lowest state of the system.

Next we find the diagonal energy $(B(\mathbf{k}-\mathbf{K},\mathbf{k})|H|B(\mathbf{k}-\mathbf{K},\mathbf{k}))$ for one of our states with a reversed spin. Most of the terms will be the same as those already considered in (12) or (13), and we shall simply consider the terms which are different. As far as the one-electron operators are concerned, we must eliminate the terms coming from $b(\mathbf{k}-\mathbf{K},\mathbf{r})$, and add extra terms coming from $b(\mathbf{k},\mathbf{r})$. It is more convenient in these expressions to retain the functions b, rather than writing in terms of atomic functions, since $\mathbf{k}$ appears explicitly in the final result. Then we find without complications that the one-electron operators give terms differing from those of (12) by

$$W(\mathbf{k})-W(\mathbf{k}-\mathbf{K})+\sum_{\mathbf{k}'}\int(b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_1)b(\mathbf{k}-\mathbf{K},\mathbf{r}_1)-b^*(\mathbf{k},\mathbf{r}_1)b(\mathbf{k},\mathbf{r}_1))b^*(\mathbf{k}',r_2)b(\mathbf{k}',\mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$-\sum_{\mathbf{k}'}\sum_{\mathbf{k}''}\int\{(b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_1)b(\mathbf{k}-\mathbf{K},\mathbf{r}_1)-b^*(\mathbf{k},\mathbf{r}_1)b(\mathbf{k},\mathbf{r}_1))b^*(\mathbf{k}',\mathbf{r}_1)b(\mathbf{k}'',\mathbf{r}_1)$$

$$\times b^*(\mathbf{k}'',\mathbf{r}_2)b(\mathbf{k}',\mathbf{r}_2)\}/\{\sum_{\mathbf{k}'''}b^*(\mathbf{k}''',\mathbf{r}_1)b(\mathbf{k}''',\mathbf{r}_1)\}2/|\mathbf{r}_{12}|d\tau_{12}. \quad (14)$$

The two-electron operators give terms differing by

$$-\sum_{\mathbf{k}'}\int(b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_1)b(\mathbf{k}-\mathbf{K},\mathbf{r}_1)-b^*(\mathbf{k},\mathbf{r}_1)b(\mathbf{k},\mathbf{r}_1))b^*(\mathbf{k}',\mathbf{r}_2)b(\mathbf{k}',\mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$-\int(b^*(\mathbf{k},\mathbf{r}_1)b(\mathbf{k},\mathbf{r}_1)b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_2)b(\mathbf{k}-\mathbf{K},\mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$+\sum_{\mathbf{k}'}\int b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_1)b(\mathbf{k}',\mathbf{r}_1)b^*(\mathbf{k}',\mathbf{r}_2)b(\mathbf{k}-\mathbf{K},\mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}. \quad (15)$$

[12] J. C. Slater, Rev. Mod. Phys. **6**, 270 (1934).

The first term represents the removal of the electrostatic interaction between the electron $\mathbf{k}-\mathbf{K}$, which is missing, with all charges, and the substitution for it of the interaction of the electron $\mathbf{k}$, which is present in its place, with all charges. The second term corrects this on account of the fact that the electron $\mathbf{k}$ must not interact with the electron $\mathbf{k}-\mathbf{K}$, which is missing. The third term removes the exchange interaction of electron $\mathbf{k}-\mathbf{K}$ with all charges. No corresponding term adding the exchange interaction of the electron $\mathbf{k}$ is present, for this electron has opposite spin to all other electrons. We now combine expressions (14) and (15), obtaining

$$(B(\mathbf{k}-\mathbf{K},\mathbf{k})\,|\,II\,|\,B(\mathbf{k}-\mathbf{K},\mathbf{k}))=E_0+W(\mathbf{k})-W(\mathbf{k}-\mathbf{K})+\sum_{\mathbf{k}'}\int b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_1)b(\mathbf{k}',\mathbf{r}_1)b^*(\mathbf{k}',\mathbf{r}_2)$$

$$\times b(\mathbf{k}-\mathbf{K},\mathbf{r}_2)2/|\mathbf{r}_{12}|\,d\tau_{12}-\sum_{\mathbf{k}'}\sum_{\mathbf{k}''}\int\{(b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_1)b(\mathbf{k}-\mathbf{K},\mathbf{r}_1)-b^*(\mathbf{k},\mathbf{r}_1)b(\mathbf{k},\mathbf{r}_1))$$

$$\times b^*(\mathbf{k}',\mathbf{r}_1)b(\mathbf{k}'',\mathbf{r}_1)\;b^*(\mathbf{k}'',\mathbf{r}_2)b(\mathbf{k}',\mathbf{r}_2)\}/\{\sum_{\mathbf{k}'''}b^*(\mathbf{k}''',\mathbf{r}_1)b(\mathbf{k}''',\mathbf{r}_1)\}2/|\mathbf{r}_{12}|\,d\tau_{12}$$

$$-\int b^*(\mathbf{k},\mathbf{r}_1)b(\mathbf{k},\mathbf{r}_1)b^*(\mathbf{k}-\mathbf{K},\mathbf{r}_2)b(\mathbf{k}-\mathbf{K},\mathbf{r}_2)2/|\mathbf{r}_{12}|\,d\tau_{12}. \quad (16)$$

By straightforward manipulation we can express this in terms of the atomic functions. We obtain the following result, where $\mathbf{R}_0$ is the radius vector to a typical nucleus, and $\boldsymbol{\varrho}$, $\boldsymbol{\varrho}'$, $\boldsymbol{\varrho}''$ are radius vectors from this nucleus to all others:

$$(B(\mathbf{k}-\mathbf{K},\mathbf{k})\,|\,II\,|\,B(\mathbf{k}-\mathbf{K},\mathbf{k}))=E_0+W(\mathbf{k})-W(\mathbf{k}-\mathbf{K})$$

$$+\sum_{\boldsymbol{\varrho},\,\boldsymbol{\varrho}'}\exp[i(\mathbf{k}-\mathbf{K})\cdot\boldsymbol{\varrho}]\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\boldsymbol{\varrho}-\boldsymbol{\varrho}')a^*(\mathbf{r}_2-R_0-\boldsymbol{\varrho}')a(\mathbf{r}_2-\mathbf{R}_0)2/|\mathbf{r}_{12}|\,d\tau_{12}$$

$$-\sum_{\boldsymbol{\varrho},\,\boldsymbol{\varrho}'}(\exp[i(\mathbf{k}-\mathbf{K})\cdot\boldsymbol{\varrho}]-\exp[i\mathbf{k}\cdot\boldsymbol{\varrho}])\int\{[\sum_{\mathbf{R}}a^*(\mathbf{r}_1-\mathbf{R})a(\mathbf{r}_1-\mathbf{R}-\boldsymbol{\varrho})]/[\sum_{\mathbf{R}'}a^*(\mathbf{r}_1-\mathbf{R}')a(\mathbf{r}_1-\mathbf{R}')]\}$$

$$\times a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\boldsymbol{\varrho}')a^*(\mathbf{r}_2-\mathbf{R}_0-\boldsymbol{\varrho}')a(\mathbf{r}_2-\mathbf{R}_0)2/|r_{12}|\,d\tau_{12}-(1/N)\sum_{\boldsymbol{\varrho},\,\boldsymbol{\varrho}',\,\boldsymbol{\varrho}''}\exp[i(\mathbf{k}\cdot\boldsymbol{\varrho}+(\mathbf{k}-\mathbf{K})\cdot\boldsymbol{\varrho}')]$$

$$\times\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\boldsymbol{\varrho})a^*(\mathbf{r}_2-\mathbf{R}_0-\boldsymbol{\varrho}'')a(\mathbf{r}_2-\mathbf{R}_0-\boldsymbol{\varrho}'-\boldsymbol{\varrho}'')2/|\mathbf{r}_{12}|\,d\tau_{12}. \quad (17)$$

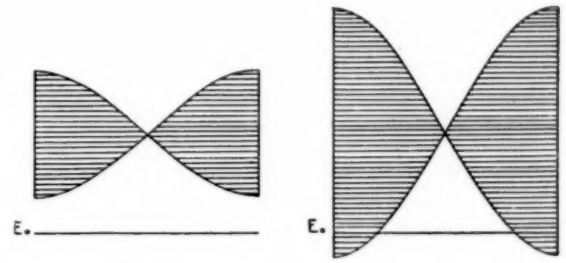
It is worth while looking a little at expression (17), for it indicates the energy as given by the energy band theory, without the application of perturbations to it. The quantity $W(\mathbf{k})-W(\mathbf{k}-\mathbf{K})$ measures the energy, as given by the energy band theory, which is necessary to take an electron from state $\mathbf{k}-\mathbf{K}$, and bring it into state $\mathbf{k}$. This of course is zero if $\mathbf{K}=0$, for then, though the spin is reversed, the electron is left in the same place in the energy band that it was before. For a small $\mathbf{K}$, the term is small, measuring the difference of energy between points separated by the vector $\mathbf{K}$ in the Brillouin zone, and can be either positive or negative depending on $\mathbf{k}$. The maximum value of the energy difference comes when $\mathbf{K}$ is the vector drawn from the point of maximum energy in the zone to the point of minimum energy, or *vice versa;* that is, ordinarily, from the center of the zone to a point on the boundary. For each value of $\mathbf{K}$, there are N states, coming from the N values of $\mathbf{k}$, and at this extreme point the N states are spread into a band whose width is twice the width of the energy band $W(\mathbf{k})$. This is indicated for a one-dimensional schematic model in Fig. 1. There we plot the band of excited levels, as a function of $\mathbf{K}$, showing that the band shrinks to zero for $\mathbf{K}=0$, increases in

width proportionally to $\mathbf{K}$ for small $\mathbf{K}$, then approaches a maximum for $\mathbf{K}$ equal to half the width of the zone (that is, the vector from the center to the edge of the zone). The whole diagram of course is periodic in $\mathbf{K}$. Similarly in three dimensions it will be periodic with the periodicity of the reciprocal lattice.

Next we must consider the summation terms in (17). The leading term comes from the case $\varrho=\varrho'=0$ in the first summation, and is

$$\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0) \times a^*(\mathbf{r}_2-\mathbf{R}_0)a(\mathbf{r}_2-\mathbf{R}_0)2/|\mathbf{r}_{12}|d\tau_{12}. \quad (18)$$

This measures the interaction of an electron on the typical atom with another electron on the same atom; that is, it measures essentially the energy required to take an electron from a neutral atom and put it on another atom which is already neutral, producing a negative ion, and leaving a positive ion behind. This term persists to infinite separation. Similarly the last summation of (17), for the case $\varrho=\varrho'=\varrho''=0$, gives the same term, but multiplied by $-1/N$. The corresponding term of the second summation of (17) is zero on account of the cancellation of the exponential functions. All other terms of the summations in (17) involve integrals between electrons on different atoms, and hence are smaller, vanishing at infinite separation. While they cannot be neglected for all purposes, as we shall see later, we may leave them out of account in a first rough survey. We see then that added to the energy $W(\mathbf{k})-W(\mathbf{k}-\mathbf{K})$ there is a term, approximately independent of $\mathbf{K}$, and persisting to infinite separation, measuring $(1-1/N)$ times the work necessary to form a positive and a negative ion. The reason for this term is clear. The excited wave functions, as given by the energy band theory, correspond to an electron of negative spin traveling through the metal. The Pauli principle does not prevent it from being on the same atom with an electron of positive spin. Then it will have only one chance in N of being on the positive ion, and $(1-1/N)$ chances of being on a neutral atom, so that its average energy will include a term $(1-1/N)$ times the interaction energy with another electron on the same atom. This is a correct average result, even for infinite separation, but it is the lowest state that interests us more than the average. And the lowest state, at infinite separation, is surely that in which the electron merely



a b

FIG. 1. Energy continuum, with one excited electron, on the energy band theory, as function of K. E_0 represents the energy of the state with all spins parallel, and the hatched areas represent the continua of excited levels. a, ferromagnetic case; b, nonferromagnetic case.

reverses its spin, without moving to another atom. This state has the same energy E_0 as the initial state of maximum spin. We see therefore that while the energy band theory may be correct, and in fact proves to be correct, for the great majority of excited states, there are nevertheless a few discrete states below the continuum, which naturally are of great importance at low temperatures. We shall see later how to find more information about these discrete states.

From Fig. 1, even without a study of the discrete states, we can obtain an important criterion for the existence of ferromagnetism. As far as our simple model goes, it is certainly reasonable to say that if all states with one negative spin have higher energy than the state with maximum positive spin, the system will be ferromagnetic, while if some states of negative spin have less energy than the state of maximum spin, so that the energy is reduced with reduction of net magnetic moment, the system will not be ferromagnetic. Thus a necessary and sufficient condition for ferromagnetism is that all the excited levels lie above E_0. Now the levels of Fig. 1 represent merely the average energy over unperturbed wave functions. The variation principle tells us, however, that when we find the correct wave functions and energy levels, the lowest correct energy levels will lie below the lowest unperturbed ones. As a matter of

fact, there will be a set of discrete levels splitting off the bottom of the continuum, and they will necessarily lie lower than the bottom of the continuum of Fig. 1. Thus we have a necessary condition for ferromagnetism, in terms of the energy band theory: the energy bands of Fig. 1 must lie entirely above E_0, or the splitting of the energy band must be less than the interaction integral between two electrons on the same atom. This is essentially equivalent to the criterion of the references 5, though that was expressed in terms of the comparison of the state of maximum spin with that of zero spin. But this condition by itself is not sufficient. In Fig. 1a we show a case where the splitting is less than the interaction integral, so that we may have ferromagnetism; but it is still possible that some discrete states should lie below E_0, in which case the system would not show ferromagnetism. Fig. 1b, on the other hand, in which the splitting is greater than the interaction integral, represents a state which cannot possibly be ferromagnetic. In order to find more definite information about the sufficient condition for ferromagnetism, we must inquire about the discrete states.

As a next step, we shall find the energy of the excitation states $F(\mathbf{K}, \mathbf{R})$. To do this, we first need not only the diagonal but also the nondiagonal energy components in the B system. It is readily proved that there is no component of energy between two states differing in $\mathbf{K}$. Any change in the $\mathbf{k}$ value of one electron must change $\mathbf{K}$, so that the only nonvanishing components come from the simultaneous change in $\mathbf{k}$ of two electrons. We shall then compute $(B(\mathbf{k}-\mathbf{K}, \mathbf{k})|H|B(\mathbf{k}'-\mathbf{K}, \mathbf{k}'))$, in which originally there is an electron of $+$ spin missing from state $\mathbf{k}-\mathbf{K}$, one of $-$ spin present in $\mathbf{k}$, and finally there is one of $+$ spin missing from $\mathbf{k}'-\mathbf{K}$, one of $-$ spin present in $\mathbf{k}'$. Stated otherwise, the original state has an electron of $-$ spin in state $\mathbf{k}$, and one of $+$ spin in state $\mathbf{k}'-\mathbf{K}$, which after transition change to an electron of $+$ spin in state $\mathbf{k}-\mathbf{K}$, and an electron of $-$ spin in state $\mathbf{k}'$. The nondiagonal component is then given by

$$(B(\mathbf{k}-\mathbf{K}, \mathbf{k})|H|B(\mathbf{k}'-\mathbf{K}, \mathbf{k}'))$$

$$=\tfrac{1}{2}\sum_{\text{spin}}\int\begin{vmatrix} b_-^*(\mathbf{k}, \mathbf{r}_1) & b_-^*(\mathbf{k}, \mathbf{r}_2) \\ b_+^*(\mathbf{k}'-\mathbf{K}, \mathbf{r}_1) & b_+^*(\mathbf{k}'-\mathbf{K}, \mathbf{r}_2)\end{vmatrix} 2/|\mathbf{r}_{12}| \begin{vmatrix} b_+(\mathbf{k}-\mathbf{K}, \mathbf{r}_1) & b_+(\mathbf{k}-\mathbf{K}, \mathbf{r}_2) \\ b_-(\mathbf{k}', \mathbf{r}_1) & b_-(\mathbf{k}', \mathbf{r}_2)\end{vmatrix} d\tau_{12}$$

$$=-\int b^*(\mathbf{k}, \mathbf{r}_1)b(\mathbf{k}', \mathbf{r}_1)b^*(\mathbf{k}'-\mathbf{K}, \mathbf{r}_2)b(\mathbf{k}-\mathbf{K}, \mathbf{r}_2)2/|\mathbf{r}_{12}|d\tau_{12}. \quad (19)$$

Now that we have the diagonal and nondiagonal matrix components of energy in the B system, in Eqs. (16) or (17), and (19), we can readily find the components in the F system, using the transformation equation

$$(F(\mathbf{K}, \mathbf{R})|H|F(\mathbf{K}, \mathbf{R}'))=\sum_{\mathbf{k}}\sum_{\mathbf{k}'}(1/N)\exp[i(\mathbf{k}'\cdot\mathbf{R}'-\mathbf{k}\cdot\mathbf{R})](B(\mathbf{k}-\mathbf{K}, \mathbf{k})|H|B(\mathbf{k}'-\mathbf{K}, \mathbf{k}')). \quad (20)$$

We find

$$(F(\mathbf{K}, \mathbf{R})|H|F(\mathbf{K}, \mathbf{R}'))=E_0\delta(\mathbf{R}, \mathbf{R}')+1/N\sum_{\mathbf{k}}\exp[i\mathbf{k}\cdot(\mathbf{R}'-\mathbf{R})](W(\mathbf{k})-W(\mathbf{k}-\mathbf{K}))$$

$$+\sum_{\rho}\exp[i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})]\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0+\mathbf{R}'-\mathbf{R}-\rho)a^*(\mathbf{r}_2-\mathbf{R}_0-\rho)a(\mathbf{r}_2-\mathbf{R}_0)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$-\sum_{\rho}\exp[i\mathbf{K}\cdot\rho]\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0+\mathbf{R}'-\mathbf{R}-\rho)a^*(\mathbf{r}_2-\mathbf{R}_0-\mathbf{R}-\rho)a(\mathbf{r}_2-\mathbf{R}_0-\mathbf{R})2/|\mathbf{r}_{12}|d\tau_{12}$$

$$-\sum_{\rho}(\exp[i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})]-1)\int\{[\sum_{\mathbf{R}_1}a^*(\mathbf{r}_1-\mathbf{R}_1)a(\mathbf{r}_1-\mathbf{R}_1+\mathbf{R}'-\mathbf{R})]/[\sum_{\mathbf{R}_2}a^*(\mathbf{r}_1-\mathbf{R}_2)a(\mathbf{r}_1-\mathbf{R}_2)]\}$$

$$\times a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\rho)a^*(\mathbf{r}_2-\mathbf{R}_0-\rho)a(\mathbf{r}_2-\mathbf{R}_0)2/|\mathbf{r}_{12}|d\tau_{12}. \quad (21)$$

This formula includes both diagonal and nondiagonal terms. One simplification can be made, in the second term, involving the W's. By the general theory of Brillouin zones, $W(\mathbf{k})$ is a periodic function in $\mathbf{k}$ space, and hence can be expanded in Fourier series:

$$W(\mathbf{k})=\sum_{\mathbf{R}} W_R \exp\left[-i\mathbf{k}\cdot\mathbf{R}\right], \tag{22}$$

where the summation is over all vectors $\mathbf{R}$ from a typical atom to all other atoms. Using this expansion, we find easily

$$1/N\sum_{\mathbf{k}} \exp\left[i\mathbf{k}\cdot(\mathbf{R}'-\mathbf{R})\right](W(\mathbf{k})-W(\mathbf{k}-\mathbf{K}))=W_{\mathbf{R}'-\mathbf{R}}(1-\exp\left[i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})\right]). \tag{23}$$

Just as with the energy band functions, it is interesting to consider the diagonal energy in terms of the F functions. Since the formulas simplify considerably when $\mathbf{R}'-\mathbf{R}=0$, it is worth while writing these diagonal terms:

$$(F(\mathbf{K},\mathbf{R})\,|H|\,F(\mathbf{K},\mathbf{R}))=E_0+\sum_{\boldsymbol{\rho}}\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\boldsymbol{\rho})a^*(\mathbf{r}_2-\mathbf{R}_0-\boldsymbol{\rho})a(\mathbf{r}_2-\mathbf{R}_0)2/|\mathbf{r}_{12}|d\tau_{12}$$

$$-\sum_{\boldsymbol{\rho}}\exp\left[i\mathbf{K}\cdot\boldsymbol{\rho}\right]\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\boldsymbol{\rho})a^*(\mathbf{r}_2-\mathbf{R}_0-\mathbf{R}-\boldsymbol{\rho})a(\mathbf{r}_2-\mathbf{R}_0-\mathbf{R})2/|\mathbf{r}_{12}|d\tau_{12}. \tag{24}$$

The leading terms come from the case $\boldsymbol{\rho}=0$. Then the first summation represents the potential of an electron on the atom at $\mathbf{R}_0$, on itself, while the second summation is the potential of an electron on R_0, on another electron on the atom at $\mathbf{r}_0+\mathbf{R}$. If $\mathbf{R}=0$, these terms just cancel, so that the energy is E_0, except for exchange terms with neighboring atoms, which we shall consider in a moment. If $\mathbf{R}\neq 0$, on the other hand, the terms do not cancel, but instead the second reduces approximately to the potential between a spherical charge on the atom at $\mathbf{R}_0$, and a spherical charge on the atom at $\mathbf{R}_0+\mathbf{R}$, which is $2/|\mathbf{R}|$. That is, the whole thing reduces to the energy necessary to form a positive and negative ion at distance $\mathbf{R}$, the Coulomb term $-2/|\mathbf{R}|$ arising from the attraction of the ions for each other. To this approximation, we shall have a series of N terms for each $\mathbf{K}$ value, the lowest one being approximately E_0, the others converging, as $\mathbf{R}$ becomes large, to a limit. Since most of the terms lie arbitrarily close to the limit, the average energy of the terms is closely equal to the limit, and by general principles of wave mechanics this average value must also equal the average value of the terms as computed in Fig. 1 from the energy band theory.

We shall now consider these states in more detail. The lowest state is for $\mathbf{R}=0$, and for it we have

$$(F(\mathbf{K},0)\,|H|\,F(\mathbf{K},0))=E_0$$

$$+\sum_{\boldsymbol{\rho}}(1-[\exp i\mathbf{K}\cdot\boldsymbol{\rho}])\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\boldsymbol{\rho})$$

$$\times a^*(\mathbf{r}_2-\mathbf{R}_0-\boldsymbol{\rho})a(\mathbf{r}_2-\mathbf{R}_0)2/|\mathbf{r}_{12}|d\tau_{12}. \tag{25}$$

The integral in (25) is the interaction of a charge distributed on atoms $\mathbf{R}_0$ and $\mathbf{R}_0+\boldsymbol{\rho}$, on itself, and hence is necessarily positive. Furthermore it is equal for all neighbors $\boldsymbol{\rho}$ at a fixed distance from $\mathbf{R}_0$, and falls off rapidly as $\boldsymbol{\rho}$ increases, so that for approximate purposes we need consider only nearest neighbors. The factor $(1-\exp i\mathbf{K}\cdot\boldsymbol{\rho})$ is zero for $\mathbf{K}=0$, showing that the energy of this state is exactly E_0 at the center of the zone, but is positive for other $\mathbf{K}$'s, showing that the energy increases for other $\mathbf{K}$ values, reaching a maximum at the edge of the zone. This behavior is shown in Fig. 2, where these energy levels are

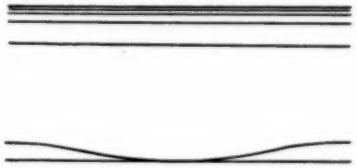


FIG. 2. Energy levels with one excited electron, on the spin wave theory, as function of K.

plotted as functions of $\mathbf{K}$, in the same way that the energy band levels were plotted in Fig. 1. This variation with $\mathbf{K}$ is just what Bloch found for his spin waves, which are essentially equivalent to our functions $F(\mathbf{K}, 0)$, with one exception: the exchange integral which Bloch had was the Heisenberg exchange integral, involving both a positive term, coming from interaction of the electrons with each other, and a negative term, coming from interaction with the nuclei. Thus if the positive term were larger than the negative, the energy of the spin wave would be larger than E_0 for all $\mathbf{K}$'s except $\mathbf{K}=0$, leading to ferromagnetism, while if the negative term outweighed the positive one, the energy would be less than E_0, and there would be no ferromagnetism. We have at this point only the equivalent of his positive term, and shall find that the negative contribution is supplied later when we perform our perturbation calculation. As for the states with $\mathbf{R}\neq 0$, the variation with $\mathbf{K}$ is unimportant. It comes from the last summation in Eq. (24). This represents the interaction between a charge shared on atoms $\mathbf{R}_0$ and $\mathbf{R}_0+\boldsymbol{\rho}$, on a similar charge displaced a distance $\mathbf{R}$. Now on account of the orthogonality of the a's, the net charge in one of these distributions must be zero, so that it is a multipole. On account of symmetry it cannot be a dipole, and must be at the least a quadripole. Thus we have the potential of a quadripole on another, falling off as a high inverse power of the distance, and becoming practically negligible after the distance $\mathbf{R}$ becomes appreciable. Thus the higher states in Fig. 2 are drawn independent of $\mathbf{K}$.

The Perturbation Problem

We have now derived the matrix components of energy in two systems, and we can proceed from either one to set up the perturbation problem to get the actual states of the system. If we start with the energy band picture, as in Fig. 1, we shall find that the perturbations split certain discrete levels from the continuum, leaving the major part of the continuum essentially unchanged. If we start on the other hand from the functions of Fig. 2, where the lowest state represents the spin waves, and the higher states represent polar spin waves, the perturbation will spread out the large number of levels near the series limit into a band or continuum, at the same time distorting the lower levels, and in general depressing them. The result of either sort of calculation is the same, and is indicated in Fig. 3, where in 3a we see the case where all of the final states lie above E_0, so that we have ferromagnetism, while in 3b some of the levels lie below E_0, so that the energy of the system can be decreased by reducing the net magnetic moment, and we do not have ferromagnetism. In either case it is the perturbed spin wave level, the one corresponding to $F(\mathbf{K}, 0)$, which results in the lowest level. If this lies above E_0 for all $\mathbf{K}$'s, we have ferromagnetism, and otherwise not. We shall therefore study this state in detail.

It is easiest to study this spin wave in terms of the functions F, using the matrix components in (21). We first consider the point $\mathbf{K}=0$. The unperturbed energy for this value of $\mathbf{K}$ is just E_0, as we saw from Eq. (25). Furthermore, an examination of (21) shows at once that the nondiagonal components from this state to other states, $(F(0, 0)\,|H|\,F(0, \mathbf{R}'))$, are all zero. Thus the function $F(0, 0)$, with the energy E_0, is an exact solution of the problem, and the lowest energy level must go through the point $\mathbf{K}=0$, energy $=E_0$, possessing a minimum or maximum there by symmetry. As a matter of fact, we can show by a perfectly general argument that there must be a state of the system with these values of $\mathbf{K}$ and energy. For out of all the N^2 states corresponding to a component of spin along the axis connected with $(N-1)$ parallel spins, 1 antiparallel spin, there must be (N^2-1) states corresponding to the same total spin, and 1 state corresponding to the particular orientation of the state of maximum spin (N parallel spins) which has a component corresponding to $N-1$ spins along the axis. This one state must, by the general theory of multiplets, have the same energy E_0 and the same value $\mathbf{K}=0$ characteristic of the state of maximum spin, which we are using as our zero of reference. Thus this level with $E=E_0$ is not really a level of the problem we are considering at all, and we could have predicted its existence from the outset. We could not, however, have predicted on general grounds that other levels should join smoothly onto this level.

We next consider the nondiagonal matrix components $(F(\mathbf{K}, 0)|H|F(\mathbf{K}, \mathbf{R}))$ between the lowest state and other states, when $\mathbf{K}\neq 0$. The leading term is the one given in (23), equal to $W_R(1-\exp i\mathbf{K}\cdot\mathbf{R})$. The other terms, in the summations, are all exchange terms between neighbors, and so are small. Furthermore, even the leading term in the summations cancels. This is the term $\boldsymbol{\rho}=0$. In the last summation of (21), in this case, the most important contributions to the summations over $\mathbf{R}_1$ and $\mathbf{R}_2$ come by setting $\mathbf{R}_1=\mathbf{R}_2=\mathbf{R}_0$. If this is done, and $\boldsymbol{\rho}=0$, it is easy to see that the summations vanish. The remaining terms are small enough so that we can safely neglect them. Thus we are left only with the nondiagonal term $W_R(1-\exp i\mathbf{K}\cdot\mathbf{R})$. We may expect that W_R is largest for interactions with nearest neighbors, $\mathbf{R}=$ distance from an atom to its nearest neighbor. In fact, if $W(\mathbf{k})$, in (22), is given by the formula of Bloch's approximate theory, the interaction W_R is zero except for nearest neighbors, negative for nearest neighbors. Thus we may expect to get a good approximation if we consider only interactions of $F(\mathbf{K}, 0)$ with those functions in which the electron of negative spin is removed to a nearest neighbor of the atom from which it has come.

It is not practical to try to get exact solutions for the lowest state of the real system, so we shall first proceed by perturbation theory. Using second-order perturbations, and considering interactions with nearest neighbors, as we have just stated, we then have

$$\text{energy}=(F(\mathbf{K}, 0)|H|F(\mathbf{K}, 0))$$
$$-\sum_{\mathbf{R}} W_R{}^2\frac{(1-\exp[i\mathbf{K}\cdot\mathbf{R}])(1-\exp[-i\mathbf{K}\cdot\mathbf{R}])}{E(\mathbf{R})-E(0)} \quad (26)$$

Here $E(\mathbf{R})$ is the unperturbed energy of the $\mathbf{R}$ state, $(F(\mathbf{K}, \mathbf{R})|H|F(\mathbf{K}, \mathbf{R}))$, and $E(0)$ is the energy of the 0 state, $(F(\mathbf{K}, 0)|H|F(\mathbf{K}, 0))$. Since $E(\mathbf{R})$, W_R are the same for all nearest neighbors $\mathbf{R}$, we can take them out of the summation. We shall adopt two abbreviations: in (25), we let the integral be I_0:

$$I_0=\int a^*(\mathbf{r}_1-\mathbf{R}_0)a(\mathbf{r}_1-\mathbf{R}_0-\boldsymbol{\rho})a^*(\mathbf{r}_2-\mathbf{R}_0-\boldsymbol{\rho})$$
$$\times a(\mathbf{r}_2-\mathbf{R}_0)2/|\mathbf{r}_{12}|d\tau_{12}, \quad (27)$$

where $\boldsymbol{\rho}$ is the vector to one of the nearest neighbors. Secondly, we let

$$I_1=E(\mathbf{R})-E(0). \quad (28)$$

Furthermore, we shall neglect the fact that I_1 depends on $\mathbf{K}$, since the dependence, which comes from $E(0)$, is not very great. We remember that

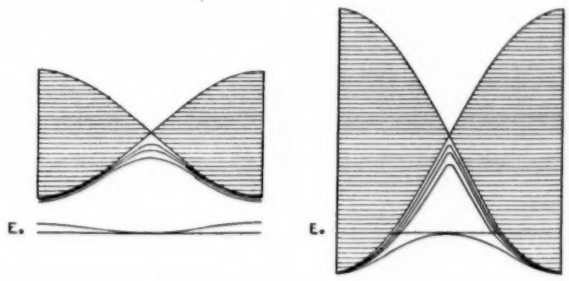


FIG. 3. Energy levels with one excited electron, as function of K, as result of perturbation theory applied to methods of Fig. 1 or Fig. 2. a, ferromagnetic case; b, nonferromagnetic case. Note: Figs. 1, 2, 3 are merely schematic, not the result of actual computations.

the largest term in I_1 is the energy of interaction between two electrons on the same atom, a fairly large quantity compared with the variation of $E(0)$ with $\mathbf{K}$. Then in evaluating (26), we must first find the summation $\sum_{\rho}(1-\exp i\mathbf{K}\cdot\boldsymbol{\rho})$ from (25). The exponential can be written as $\cos\mathbf{K}\cdot\boldsymbol{\rho}+i\sin\mathbf{K}\cdot\boldsymbol{\rho}$, of which the second term is odd in $\boldsymbol{\rho}$. Now for each neighbor there will be another diametrically opposite, so that the sine terms will cancel in pairs, leaving only the cosine terms. Hence

$$\sum_{\rho}(1-\exp[i\mathbf{K}\cdot\boldsymbol{\rho}])=\sum_{\rho}(1-\cos\mathbf{K}\cdot\boldsymbol{\rho}). \quad (29)$$

Also we must work out the sum in (26). This is

$$\sum_{\mathbf{R}}(1-\exp[i\mathbf{K}\cdot\mathbf{R}])(1-\exp[-i\mathbf{K}\cdot\mathbf{R}])$$
$$=\sum_{\mathbf{R}}2(1-\cos\mathbf{K}\cdot\mathbf{R}), \quad (30)$$

twice the previous summation. Thus we have

$$\text{energy of lowest state}=E_0+A\sum_{\mathbf{R}}(1-\cos\mathbf{K}\cdot\mathbf{R}),$$

where

$$A=I_0-2W_R{}^2/I_1. \quad (31)$$

The formula (31) gives a variation with $\mathbf{K}$ just like Bloch's solution for the same spin wave, as we have mentioned above. Now, however,

we get more complete agreement with the Bloch or Heisenberg type of theory, in which A would take the place of the interaction or exchange integral between neighbors. For now our integral A is made up of two parts, a positive and a negative one, the positive one coming from electrostatic interactions between electrons, the negative one from the splitting of the energy band, which amounts to interactions between electrons and nuclei. Plainly the quantity A can be of either sign, and if it is positive we have ferromagnetism, otherwise not. (We must remember that our formula, being derived from perturbation theory, is not very accurate for large perturbations, or large W_R. Thus it should be fairly good for the ferromagnetic case, but inaccurate for the nonferromagnetic case.) It is interesting to consider how A will change as the energy band broadens. Both I_0 and W_R depend on the overlapping of neighboring atoms, though in a rather independent way. For purposes of orientation, we may suppose that they are proportional to each other: $I_0=\alpha|W_R|$, where probably α is considerably smaller than unity. Then we have

$$A=\alpha|W_R|-2W_R^2/I_1, \tag{32}$$

a quadratic in W_R, going to zero when $W_R=0$, rising to a maximum as the overlapping increases, then going to zero again when $|W_R|=\alpha I_1/2$, finally becoming and staying negative. This is the behavior which has often been discussed, in which ferromagnetism starts weakly with small overlapping of neighbors (as in the rare earths), increases with greater overlapping (as in the iron group), then disappears with still more overlapping (as in the conduction electrons of nonferromagnetic elements).

A more accurate solution of our perturbation problem can be obtained in some special cases. Thus consider a body centered lattice, and let $\mathbf{K}$ be at the edge of the band. That is, $\mathbf{K}$ is along one of the three axes, and its magnitude is such that $\cos \mathbf{K}\cdot\mathbf{R}=-1$ for each of the eight neighbors. Our problem is then to find combinations of all functions $F(\mathbf{K}, \mathbf{R})$, for all $\mathbf{R}$'s, which will be solutions of the problem. Remembering that each atom has eight neighbors, the diagonal energy of the state $F(\mathbf{K}, 0)$ is E_0+16I_0. To the approximation which we have been making, we can set the diagonal energy of all other states $F(\mathbf{K}, \mathbf{R})$ for $\mathbf{R}\neq 0$ equal to E_0+I_1. Furthermore, the nondiagonal matrix component from one function to any one of its eight nearest neighbors is $2W_R$, and to any further neighbor we assume it to be zero. By symmetry, the lowest state will then have a wave function in which all functions with the same magnitude of $\mathbf{R}$, but different directions, will have the same coefficient. Thus let the function $F(\mathbf{K}, 0)$, the nonpolar function, have the coefficient C_0. The next more distant functions, corresponding to the removal of the electron to one of the eight nearest neighbors, at $(\pm 1, \pm 1, \pm 1)$ times the lattice spacing, will have the coefficient C_1. The six functions connected with the next nearest neighbors $(\pm 2, 0, 0)$, $(0, \pm 2, 0)$, $(0, 0, \pm 2)$, will have the coefficient C_2. We continue in this way. Then it is easy to set up the secular equation for the energy. This is found, by consideration of the number of neighbors of each type which each function has, to be the following:

$$\begin{vmatrix} E_0+16I_0-E & 16W_R & 0 & 0 & 0 \\ 2W_R & E_0+I_1-E & 6W_R & 6W_R & 0 \\ 0 & 8W_R & E_0+I_1-E & 0 & 8W_R \\ 0 & 4W_R & 0 & E_0+I_1-E & 4W_R \\ 0 & 0 & 2W_R & 4W_R & E_0+I_1-E \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{vmatrix} =0. \tag{33}$$

We can approximate this successively by a quadratic (using the four elements in the first two rows and columns), a cubic, quartic, etc., using successively more rows and columns. This has been done, carrying it up to the fifth degree equation, which amounts to using the nonpolar

state and the four nearest sets of neighbors. In Fig. 4 the results are plotted for each of these approximations, as well as for the second order perturbation method of Eq. (31). The integral I_1 in Fig. 4 is taken as the unit of energy, and the quantity $16|W_R|$ is used as the abscissa, so that the figure shows what happens as the energy band becomes broader and broader. The straight line represents $I_1-16|W_R|$, which is easily seen to be the bottom of the Bloch continuum, in case the zero of energy represents E_0. The point where it crosses the axis is the point where the theory of energy bands would indicate that ferromagnetism ceased. The various curves, all approximations to one correct solution of the perturbation problem, indicate the lowest stationary state of the problem, essentially Bloch's spin wave, as it is determined from the secular Eq. (33). For $W_R=0$, this reduces simply to E_0+16I_0, and we have shown the case where $16I_0=1/10I_1$. That is, we have assumed the interaction integral of the electron with its neighbors ($16I_0$) to be one-tenth the interaction with a similar electron on the same atom (I_1). This seems like a reasonable assumption in the ferromagnetic case.

Still another approximation curve is shown in Fig. 4, lying lower than the others, and therefore still nearer the truth. This was obtained by a method similar to Wannier's operator method.[10] Let us suppose that the coefficients C vary slowly from one atom to its neighbors. We shall now label the C's with the indices of the atom to which they refer: $C(n_1n_2n_3)$ is the coefficient of the function $F(\mathbf{K},\mathbf{R})$ for which the vector $\mathbf{R}$ equals (n_1, n_2, n_3) times the lattice spacing. Then, so long as $n_1n_2n_3$ is neither the central atom, $\mathbf{R}=0$, nor one of its eight nearest neighbors, the corresponding Eqs. (33) all reduce to the form

$$(E_0+I_1-E)C(n_1n_2n_3)+2W_R(C(n_1+1, n_2+1, n_3+1)+\cdots+C(n_1-1, n_2-1, n_3-1))=0, \quad (34)$$

where there are eight terms multiplying $2W_R$. Now suppose that we can expand the C's in power series in the n's, and break off with second power terms, which we can do if the C's vary sufficiently slowly. The linear terms will vanish, equal and opposite contributions coming from

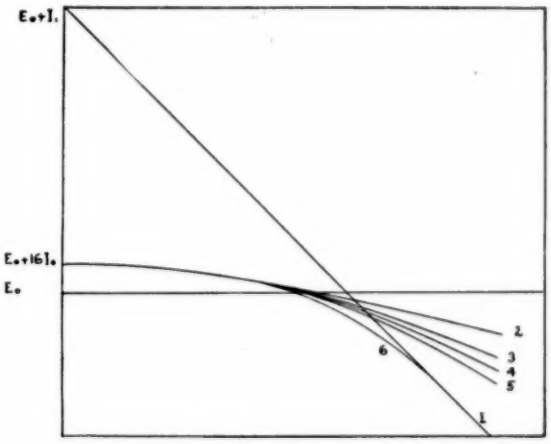


FIG. 4. Various approximations to the lowest energy level, as function of W_R, the width of the energy band. Straight line (1) represents the bottom of the energy continuum. Curves (2), (3), (4), (5), are solutions of the second, third, fourth, and fifth power secular equations approximating the Eq. (33). Curve (6) represents the solution of Eq. (36). The second order perturbation approximation of Eq. (31) lies very close to curve (2).

opposite neighbors. We are then left only with terms coming from the second derivatives. These give the equation

$$(E_0+I_1+16W_R-E)C(n_1n_2n_3)+8W_R(\partial^2/\partial n_1^2+\partial^2/\partial n_2^2+\partial^2/\partial n_3^2)C(n_1n_2n_3)=0. \quad (35)$$

Eq. (35) is the wave equation, and we find the desired spherically symmetrical solution by assuming

$$C(n_1n_2n_3)=(\exp -\alpha(n_1^2+n_2^2+n_3^2)^{\frac{1}{2}})/(n_1^2+n_2^2+n_3^2)^{\frac{1}{2}}. \quad (36)$$

where α is an arbitrary constant. Between α and the energy there must exist the relation

$$E=E_0+I_1+16W_R+8W_R\alpha^2. \quad (37)$$

Eq. (37) has a simple interpretation in terms of Fig. 4. Remembering that W_R is negative, we see that $E_0+I_1+16W_R$ represents the straight line of Fig. 4, or the bottom of the energy band continuum. Eq. (37) then says that the energy E lies below this energy, by an amount which is smaller, the smaller α becomes. Small α corresponds to a wave function which is reduced only very slowly as $(n_1^2+n_2^2+n_3^2)$ increases; that is, to a very extended wave function.

The expression (36) clearly cannot be used for $n_1=n_2=n_3=0$, for it then becomes infinite. Let

us then assume (36) for all other values of the n's, and assume an independent value C_0 for the central atom. We then have three variables, C_0, α, and E, and we must find three equations to determine them. One of these equations is (37), and it is natural to take as the other two the first two equations of those indicated in the determinant (33), so that the relations will be exactly satisfied for the central atom and its next nearest neighbors. Eliminating C_0, these give a relation between the energy and α which can be easily solved, and the result is plotted in Fig. 4. As we see, the energy of the state gradually approaches the bottom of the continuum, with corresponding decrease of α, or increase of the size of the wave functions. Even this solution proves not to be good for very large values of W_R; the solution shows a curious behavior by which α becomes zero for a finite value of W_R, which surely does not correspond to the real behavior. It is safe to say, however, that this solution is fairly near the truth in the region where it lies above the energy E_0, which is the ferromagnetic case, and that it indicates the correct trend even when it lies below E_0.

The situation, then, as indicated by Fig. 4 and the calculations leading up to it, is the following. The energy band of Figs. 1 and 3, as it increases in width, pushes the lowest wave function down, and it pushes it below E_0, destroying ferromagnetism, when the energy band is about eighty percent of the integral I_1, rather than one hundred percent as in the elementary energy band theory. This numerical value, of course, represents a special case, but it probably indicates the order of magnitude to be expected in general. Until the band is almost of this width, the simple second order perturbation calculation of Eq. (31) represents the energy quite accurately, and can be used safely in the ferromagnetic case. The wave function in this case consists of a mixture of the function $F(\mathbf{K}, 0)$ and the functions $F(\mathbf{K}, \mathbf{R})$ in which $\mathbf{R}$ is the nearest neighbor to the central atom. We remember that $F(\mathbf{K}, 0)$ is a function in which an electron is removed from a certain atomic function, its spin is reversed, and then it is replaced in the same atomic function. Finally all such functions are combined to make a wave of spin traversing the metal. But the atomic function, being one of the orthogonal functions, is not really confined entirely to one atom, but has appreciable values on nearest neighbors, so $F(\mathbf{K}, 0)$ contains an appreciable probability that the reversed electron will be found on a nearest neighbor. $F(\mathbf{K}, \mathbf{R})$ consists largely of the probability that the reversed electron will be on the nearest neighbor, and it enters with such a sign as partly to cancel the probability of being on the neighbor already present in $F(\mathbf{K}, 0)$. In other words, the final wave function is one in which the reversed electron is almost sure to be on the atom from which it originally came, rather than forming an ion by shifting to its neighbor. This is essentially the wave function of the Heisenberg theory, made of nonorthogonal functions, though doubtless there will be some small amount of polar function left in it. Our final result, then, is in close agreement with the atomic type of theory, as far as the lowest stationary state in the ferromagnetic case is concerned.

In the other limit, however, when the width of the energy band, W_R, becomes large, our solution departs widely from the atomic type. In the first place, the energy levels of the discrete states draw in very closely to the bottom of the continuum, so that the energy band method becomes very accurate for a calculation of energy. In the second place, as we see from the type of solution (36), the wave function changes its character entirely. It is a combination of many $F(\mathbf{K}, \mathbf{R})$'s, corresponding to large $\mathbf{R}$ values, with coefficients which fall off slowly with the magnitude of $\mathbf{R}$. That is, it is a wave function in which the electron is only tied loosely to the positive ion which it has left behind, and wanders at considerable distances from it through the crystal, though not reaching infinite distance. Our solution unfortunately is not accurate enough to give exact values for the size of this wave function. In a qualitative way, however, it is similar to those found by Wannier[10] in a similar problem. The equations (34) of course have other, higher solutions, in addition to the one we have found, which would be analogous to the 1s state of a hydrogen atom. There will be a series of s-like states, in which all neighbors at a given distance have the same coefficients, and in which the C's can be determined approximately from the differential equation (35).

Then there will be solutions of different symmetry, corresponding to p, d, $\cdots$ states of the hydrogen atom. These will form a series of discrete levels, which may have a larger or smaller total number of levels according to circumstances, lying between the lowest level, which we have investigated, and the continuum. And the size of the wave function will grow larger and larger, as we go to higher and higher "quantum numbers," regarding the wave function as sort of an "atom" on a large scale. It is this type of solution which is of importance in nonferromagnetic substances, and which we discuss in connection with the problem of superconductivity in the following paper.[13]

Conclusion

Starting with the theory of energy bands, we have set up the perturbation problem and solved approximately the case of a band containing half enough electrons to fill it, all having parallel spins but one. This problem is a test for ferromagnetism: if the lowest energy of the problem is lower than the energy when all have parallel spins, the system will tend to reduce its spin, and will not be ferromagnetic, whereas if all energies of the problem are higher than the state of maximum spin, the latter will be the stable state, and we shall have ferromagnetism. On the theory of energy bands, applied in an elementary way, as in reference 5 on the ferromagnetism of nickel, the lowest state of this problem of $(N-1)$ positive spins is found by taking the electron of positive spin from the top of the band, introducing it with negative spin at the bottom of the band, thereby reducing the energy a good deal. This reduction of energy is partly counterbalanced by the exchange integral, an interaction between an electron and another on the same atom, which increases the energy when the spin is reversed. If the energy reduction on account of the energy band is less than the exchange integral, there will be ferromagnetism, otherwise not. When we carry out our present more refined calculation, we find that in broad outlines this criterion is still a correct one. We find, however, that there are discrete energy levels below the continuum given by the energy band theory, and that some of these may lie below the energy of the state of maximum spin, destroying ferromagnetism, even when the elementary theory would predict ferromagnetism. These discrete states are states in which the electron of negative spin stays on or in the immediate neighborhood of the atom to which it formerly was attached with positive spin. When it had positive spin, the Pauli principle prevented it from passing to neighboring atoms, since they already had similar electrons. This prohibition is removed when the spin reverses, and the only thing holding the electron to its atom is now electrostatic action, or the correlation energy between the electron and positive ion. The energy by which the lowest discrete level lies below the continuum is a measure of the correlation energy. When we consider this correlation energy, we find that ferromagnetism will be destroyed when the reduction of energy on account of the width of the energy band is something like eighty percent of the exchange integral, the difference being made up by the correlation energy. Thus the criterion is changed quantitatively, but not qualitatively.

The lowest energy level of the system, the correlation state just described, is equivalent to the spin wave of Bloch's theory, and makes direct connection with Heisenberg's formulation of ferromagnetism. But we have described it in terms of orthogonal functions and energy bands, gaining in mathematical simplicity by the elimination of nonorthogonality integrals, and making connection with other parts of the theory of metals. At low temperatures, and perhaps at ordinary temperatures, it will be this spin wave which is excited, and probably the higher waves in the continuum will not be excited. Thus Bloch's deduction of the temperature variation of the magnetic moment at low temperatures, based on spin waves, should be correct, and a deduction based on the levels in the continuum is probably not valid.

In the nonferromagnetic case, where the energy bands are broad, there are still separated, discrete levels, but they approach closer and closer to the continuous band as it broadens. At the same time the wave function changes, becoming more and more extended. Instead of corresponding closely to a spin wave, the lowest state describes an electron which wanders far

[13] J. C. Slater, Phys. Rev. (following paper, this issue).

from the positive ion which it has left, describing an orbit similar to a large atom, but not going to infinite separation. This type of energy level and wave function, the energy lying only slightly below the continuum so that it will be important only at low temperature, and the wave function being very extended so that it will have large diamagnetism, may well be important in the theory of superconductivity, as discussed in the following paper.

The writer wishes to thank Professor J. H. Van Vleck for valuable discussions of the subject of this paper. He also wishes to thank the Institute for Advanced Study for the privilege of conducting these investigations in residence at Princeton.

AUGUST 1, 1937 PHYSICAL REVIEW VOLUME 52

The Nature of the Superconducting State. II

J. C. SLATER
*Institute for Advanced Study, Princeton, New Jersey**

(Received May 7, 1937)

The discrete levels of the electrons of a metal, lying below the continuum of levels predicted by the energy band theory, and interpreted in an earlier paper as leading to the superconducting state, have been further investigated, though a quantitative discussion in the general case is still impossible. The wave functions correspond to electrons which can wander for some distance through the metal, but are held to a finite region by forces of interaction with positive ions. Such wave functions will carry no current in the ordinary way, for they correspond to the correlation of an electron and a positive ion, and the two move together. On the other hand, being similar to large atoms, they have a large diamagnetism, and hence may perhaps lead to London's form of theory of superconductivity. In the second section, this possibility is discussed. It is shown, by reference to the ordinary theory of diamagnetism, that the two conventional types of theory, one for bound electrons, the other for free electrons, are treated in such different ways that one cannot in all cases interpolate between them. Instead, as wave functions become larger and larger, one can continue to treat them by the method appropriate to isolated atoms, until they become so large that the energy associated with the Larmor precession becomes comparable with the atomic energy. Then the properties change, and the method appropriate to free electrons gradually becomes correct. This limiting size depends on the magnetic field, or conversely the limiting magnetic field depends on the size. It is shown that to produce superconductivity the orbits must be of the order of magnitude of 137 atomic diameters, a not unreasonable figure with our model. Then the limiting magnetic field, above which the large diamagnetism or superconductivity would be expected to disappear, proves to be of the order of a few hundred gauss, or the order of magnitude of fields actually necessary to destroy superconductivity.

I. The Wave Functions

In an earlier paper, the author has suggested[1] that the electrons in the superconducting state may be in special stationary states of the system as a whole, lying a little below the lowest state as described by the Bloch theory of energy bands, and expressible only as a linear combination of Bloch functions, meaning that a certain correlation or cooperation between the electrons is necessary to bring this state about, which would be disturbed by temperature agitation. These special stationary states have now been further investigated, and in the present paper their nature is described in more detail, and it is shown that it is even more plausible than before that they may be really responsible for superconductivity. In the first section we discuss the nature of the wave functions and energy levels of the problem. The second section is devoted to showing that the wave functions are of the sort to be expected for superconductivity. London[2] has objected quite properly to the earlier paper, on the ground that superconductivity has much closer resemblance to diamagnetism than to ordinary conduction, a point of view which he has elaborated on previous occasions.[3] It is very

* On leave from the Massachusetts Institute of Technology, Cambridge, Mass.

[1] J. C. Slater, Phys. Rev. **51**, 195 (1937).

[2] F. London, Phys. Rev. **51**, 678 (1937).

[3] F. and H. London, Proc. Roy. Soc. **A149**, 71 (1935); Physica **2**, 341 (1935); F. London, Proc. Roy. Soc. **A152**, 24 (1935), and others. See particularly F. London, *Une Conception Nouvelle de la Supra-Conductibilité, Actualités Scientifiques et Industrielles* (Paris, 1937).

gratifying that the present theory has rather automatically developed along the lines of London's ideas, in that the special states previously described prove not to carry current in the ordinary way, but on the other hand prove to be just the sort capable of showing an extremely large diamagnetism.

It is now possible to discuss fairly completely the problem of a crystal containing two energy bands, each capable of holding N electrons, but one actually occupied by $N-1$, the other by one, electron.[4] We shall first describe the solution in nonmathematical language, hoping to understand its nature well enough so that we can indicate its extension to the actual problem in which a number of energy bands are partly occupied. This general problem is much too difficult to work with mathematically at the present time, and it is hard to see how the theory can be applied quantitatively to it without extreme complication, so that a qualitative discussion must suffice for the present.

At the outset it should be stated that we shall regard an energy band as a group of N levels, capable of holding N electrons per crystal of N atoms, of a given spin, and that we shall regard a group of N levels associated with one spin as being a different band from the group of N levels associated with the opposite spin. The essential point is that the one electron changing from one band to the other must change its atomic wave function, and it is immaterial to the general theory whether it changes the dependence of this atomic wave function on coordinates or on spin. It is for this reason that the ferromagnetic problem, where an electron reverses its spin but stays in the same band as far as orbital functions are concerned, is closely analogous to the optical absorption problem, where the electron changes its orbital function without change of spin. This observation shows that in the previous paper on superconductivity,[1] it was incorrect to suppose that the alkali metals, and copper, silver, and gold, are essentially different from other metals on account of having electrons in only one energy band. For they have electrons of both spins, and from the standpoint of the general theory that amounts to two bands, just as with other metals. In view of this fact, and of recent experimental work pointed out by Mendelssohn,[5] it seems unlikely that the theory in its present form is far enough advanced to predict which metals should show superconductivity and which should not. As we shall describe it, it certainly seems like a phenomenon which might be very widespread.

Let us call the almost filled band A, the almost empty one B. First consider the state where A is entirely filled, B entirely empty. Then the Pauli principle operates between all the N electrons in state A, in such a way as to prevent two of them from piling up on one atom at the same time. The result is an extraordinarily uniform charge distribution, quite without local fluctuations, except of course in going from one part of an atom to another. If we suppose that there are enough electrons just to neutralize the nuclei, this means a very constant electrical potential throughout the crystal. Whenever an electron moves from one atom to another, the Pauli principle demands that one move from the second atom back to the first. This shows us at once why no current can flow in such a state, since every motion of electricity is compensated by an equal and opposite motion.

Next imagine that one electron is removed from state A, placed in state B, and let us describe the situation in terms of energy bands, so that there is an empty level among the N states of band A, a filled level in band B. Of course, the lowest resulting energy, in the energy band theory, is found if the electron is removed from the highest state of band A, put in the lowest state of band B. We shall refer to the zero of energy throughout as the energy of the state where band A is filled, and we shall now consider the energy of this new state, which we shall call the excited state, referred to that zero. This excited state can be higher than the original state, in which case the system tends to have band A filled, band B empty, or it can be lower, in which case electrons tend to occupy both bands.

If the electron in band B is in a modulated wave function such as one has in a periodic lattice, that means that it is circulating freely through the crystal, and is as likely to be found

[4] J. C. Slater and W. Shockley, Phys. Rev. **50**, 705 (1936); G. Wannier, Phys. Rev. **52**, 191 (1937); J. C. Slater, Phys. Rev. **52**, 198 (1937).

[5] K. Mendelssohn, Phys. Rev. **51**, 781 (1937).

on one atom as on another. This circulation is now permitted by the Pauli principle, for this electron no longer has the same atomic wave function as the other electrons, and can occupy an atom already holding another electron. But there is a price demanded for the privilege of free circulation. This price is the energy required for the electron to be removed from its original atom (essentially the ionization potential, as distorted by the existence of the lattice), minus the energy recovered when it forms a negative ion on another atom (the electron affinity of the neutral atom, again distorted by the lattice). If the electron is not removed to a very distant atom, this is diminished by a Coulomb term: as the negative ion is removed from the positive ion left behind, work must be done against the Coulomb attraction between the ions, so that the total work which we reckoned above, representing the work necessary to remove the ions to infinite separation, must be decreased if the separation is only finite. Now the actual excited electron is equally likely to be found on any atom of the crystal, so that there is only a negligible chance that it will be near the positive ion, and this Coulomb term can ordinarily be neglected. The price demanded for the formation of the excited state then is essentially an ionization potential minus an electron affinity, a matter of several volts. But in the cases we are interested in, it is worth the price, for we shall assume that the bottom of band B lies well below the top of band A, far enough below so that the gain in energy is more than the loss, and the excited state really lies below the normal one. If the excitation consists of reversal of spin, without change of orbital motion, this means that we are not dealing with a ferromagnetic system. If it consists of change from one band to another, it means that the second band B, though it may be higher than A, still overlaps it enough so that the bottom of B is decidedly below the top of A. We find in either of these cases that the lowest state, in the language of energy bands, is that in which the excited electron is essentially at rest at the bottom of band B, the hole in band A being at the top of that band, and therefore again essentially at rest (since the slope of the energy *vs.* k curve gives the velocity, and this is zero both at the maximum and minimum energies). But these ions are both distributed with equal probability on each atom of the system, so that they have no relation to each other.

The state which we have thus derived is one in which there is a fluctuation of charge distribution. The positive and negative ions are equally likely to be found anywhere within the crystal. There is no longer the uniformity of charge and potential found in the normal state. Now a nonuniform charge distribution results in an increased potential energy. This can be seen from the fundamental principles of electrostatics. The electrostatic, potential energy can be written as $(\frac{1}{8}\pi)\int E^2 d\tau$, reducing to zero in a uniform potential where E is zero, but increasing wherever there are nonuniformities of potential. It is this increase of potential energy which we have described as the price which must be paid in order to excite an electron. We now ask, have we perhaps paid too great a price? Could we have set up some other state, with less fluctuation of charge density and potential, which would have had an even lower energy? To have less fluctuation, we must correlate the motions of positive and negative ions, arranging the state so that the chances are large that they will be found close together, small that they will be far apart. This can be done by building up a wave packet of excited states of the form we have used, each corresponding to an electron removed from one level of band A, and placed in one level of band B. If we wish to go to an extreme, we can arrange the packet so that the negative and positive ions are practically sure to coincide, and we have to pay no price at all for our excitation in the matter of potential energy. But we have a compensating loss in other energy, for it turns out that to make such a packet we must use all possible excited states, with equal coefficients, amounting to removing the electron from a mean position in band A, placing it in a mean position in band B, instead of going from the top of A to the bottom of B. In the particular case of ferromagnetism, bands A and B coincide in energy, so that in some cases it can turn out that the resulting loss of energy is very small. But if band B lies above A, the loss of energy would be considerable, and this state would probably have about as high an energy as the energy band state, or even much higher.

The situation we have just described can be looked at in a more graphic way. If the electron in band B practically coincides with the hole in band A, its wave function must decrease very rapidly as it leaves the hole. This rapid decrease, perhaps of an exponential sort, means a large kinetic energy. In other words, the increase of kinetic energy associated with such a concentrated packet more than makes up for the decrease of potential energy, in most cases. To get the really lowest energy, we must make a compromise. We build a more extended packet, so that the positive and negative ions may be found at a distance of a good many atomic diameters, but still will not wander to infinite separation from each other. This will not decrease the potential energy so much as before, but neither will it give a function that varies as much with position, so that the kinetic energy will not be as much increased. The resulting wave function can be looked at in a very physical way. Both electron and positive hole are capable of moving through the crystal (the hole "moving" by the shift of neighboring electrons into the hole, leaving a new hole on the neighboring atom). Each has a certain "effective mass," as found in the usual theory of metals, determined by the curvature of the energy-momentum curve. There is a potential energy of interaction between them. Then the wave function is approximated by that of a two-body problem with the effective masses and potential energy just described. Such a two-body problem would be similar to a hydrogen atom, except that both positive and negative particles are mobile. It has many energy levels and wave functions, the wave functions getting more and more extended as we pass to the limiting continuum at the limit of the series, and the energy levels getting closer and closer together. The levels considered in the conventional theory are simply the continuum of levels beyond the series limit, in which the electron and hole independently wander through the crystal. The discrete stationary states correspond to the discrete energy levels of a two-body problem.

This elementary picture gives one definite piece of information about these discrete wave functions: they cannot carry current in the ordinary way. For if such a function moves bodily through the crystal, both the electron and the hole move with the same rate, corresponding to equal transports of negative and positive electricity, which just cancel. Thus the existence of such levels cannot lead to a theory of enhanced ordinary conduction, as was suggested in the earlier paper.[1] In fact, if many electrons of the metal went into such states, the ordinary conductivity would be decreased, not increased.

On the other hand, there is another property of these wave functions which is just what is needed to produce a large diamagnetism. In most actual cases, the wave functions will presumably be quite extended, allowing the positive and negative charges to wander to a considerable distance from each other. That is, they will be analogous to large "atoms," much larger than ordinary atoms, but small compared to the crystal as a whole. Now the diamagnetic susceptibility of an atom is proportional to the square of its radius. Thus wave functions of this type, having a large radius, can correspond to very large diamagnetism, and can lead to London's theory of superconductivity. This is discussed in the second section.

The discussion so far has been concerned with the case in which one electron only was in band B, the remaining $N-1$ electrons being in band A. Actually, most metals have a number of bands simultaneously occupied by electrons. If the metal is not ferromagnetic, each band of one spin will be accompanied by the corresponding band of the opposite spin, and each will be equally occupied. Furthermore, practically all metals except the alkalies, copper, silver, and gold have at least two different energy bands partly occupied. The resulting problem would be extremely involved to treat in detail, and we shall not try it. But we can see qualitatively what will happen. If we set up the lowest level according to the theory of energy bands, with a number of partly occupied bands, there can be large fluctuations in the charge on each atom, for an electron corresponding to any one of the bands can be present or absent on an atom, independently of the others. Thus, as before, this state will correspond to considerable fluctuations of the potential from point to point, and to a considerable potential energy. In an effort to reduce this fluctuation, there will be special states into which each of the electrons can fall,

localized near the minima of potential energy, with energies slightly lower than those of the continuum. We can imagine a crude method of successive approximations by which these levels might be found. We could start with the energy band function. A particular, typical arrangement of electrons in it would correspond to a certain fluctuating charge distribution, with certain minima and maxima of potential. We could imagine all the electrons but one frozen in these particular positions, and let the remaining one move in the resulting field. It would have possible wave functions concentrated near the minima of potential energy, avoiding the maxima. Having found these functions, we could set up a new approximation, by putting the electrons in the lowest functions of this type, instead of in the energy band states. We should presumably find that the potential in this case had been overcorrected, and a repetition of the same process would result in more extended functions as a next approximation, and so on. Whether such an approximation would actually converge is a question, and it certainly would not be practical for computation, but it suggests a possible state of the system as a whole in which many or most of the electrons would have wave functions which were damped off in space instead of extending through the whole crystal.

This type of calculation reminds one of that made by Wigner and Seitz[6] in discussing correlation energy in a crystal. There they considered particularly the way in which the wave function of one electron kept out of the immediate neighborhood of another electron. They found that by setting up wave functions in which this was accomplished, a very considerable gain in energy can be secured. Surely this calculation of Wigner and Seitz accounts for the larger part of the gain in energy which can be secured by electron correlation, and of course this is such a large energy gain that it is not affected by temperature. The small scale fluctuations in potential will be almost entirely ironed out by this small scale correlation, and the only fluctuations which may be supposed to remain will be rather long scale fluctuations of small amplitude. These will result in only a slight excess of potential energy, so that no great energy gain will result from removing them. Nevertheless it seems reasonable to suppose that states are possible in which these large scale fluctuations are largely ironed out as well, resulting in a still further slight energy gain. In such states, the wave functions of the electrons would show large scale modulation, again resembling the large "atoms" of our earlier picture of one excited electron, though we can no longer think of a single electron and hole as being correlated, but rather of an electron being correlated with a region of slight net positive charge. To make any significant energy gain by combining such functions, we should have to arrange them all in a very accurate way, just balancing charges so as to neutralize the potential fluctuations. This would mean that such a state would be one demanding large scale cooperation of the electrons, so that it would have analogies to other cooperative phenomena such as ferromagnetism and the formation of superlattices. We should expect, then, that such a state would be realized only at low temperatures, and that there would be a sharp transition point between it and the ordinary state of the metal, as is characteristic of the transition from ordered to disordered states.

We arrive, then, at a picture of the ordinary state as one in which the electrons move more or less freely through the metal, but modifying their wave functions in the way described by Wigner and Seitz when two electrons happen to come close together, so as to remove small scale fluctuations of potential, the ones which contribute most to the excess potential energy. At lower temperatures, a different phase would come in, in which the remaining large scale fluctuations would be largely removed, by replacing the wave functions by others which, rather than extending through the crystal, are damped off in a certain finite distance, and arranging the electrons in these functions in such a way as to reduce fluctuations as far as possible. This picture is unfortunately very vague, and it seems difficult to make it less so. In particular, we cannot calculate with any degree of accuracy the energy difference between the two types of states, or the size of the wave functions in the correlated state. The energy difference certainly must be small, compared with the ordinary correlation energy,

[6] E. Wigner and F. Seitz, Phys. Rev. **46**, 509 (1934); E. Wigner, Phys. Rev. **46**, 1002 (1934).

and the wave functions large compared with the size of an ordinary atom. Our suggestion is that this correlated state corresponds to the superconducting state. If that is the case, the energy difference must be of the order of magnitude of kT for a few degrees absolute. And we shall show in the next section that the wave functions must be of the order of magnitude of several hundred atom diameters, on the average. We shall not attempt to show that the model actually leads to such figures, though of course it is to be noted that the change from wave functions modulated on a large scale, to those extending through the crystal, will be accompanied by a smaller and smaller energy change, the larger the range of the modulation is.

2. The Superconducting Property

In this section we shall largely free ourselves from the details of our model, and ask a definite question: Is it possible to have wave functions, falling off in a certain distance so as to act like big "atoms," which would show the superconducting property? At first sight, we may think of the theory of diamagnetism of single atoms, in which the diamagnetic susceptibility increases proportionally to the square of the radius. By making the radius as large as we please, we can increase the diamagnetism to any desired extent. London's[3] theory corresponds in most, though not quite all, details to the supposition that a superconductor is a substance with magnetic susceptibility $\chi=-\frac{1}{4}\pi$, so that $B=0$ inside the material. (As London points out in the references above, this is not a very accurate description, but it can serve for the moment in determining orders of magnitude.) Let us see how large the atoms would have to be to achieve this amount of diamagnetism. If there are N diamagnetic electrons per unit volume, each with a mean square radius $[r^2]_{Av}$, the theory of the diamagnetism of separate atoms gives $\chi=-e^2[r^2]_{Av}N/6mc^2$, where e, m, c have their usual meanings. We equate this figure to $-\frac{1}{4}\pi$. Further, we let $N=1/d^3$, where d is the side of a cube containing one diamagnetic electron. We shall assume this to be of the order of magnitude of interatomic distances, so that there is essentially one such electron per atom. Then an elementary calculation gives

$$([r^2]_{Av}/d^2)^{\frac{1}{2}}=137(3d/2\pi a_0)^{\frac{1}{2}}, \qquad (1)$$

where a_0 is the atomic unit of distance. If d is of the order of interatomic distances, $(3d/2\pi a_0)^{\frac{1}{2}}$ will be of the order of unity. Hence to get enough diamagnetism to agree with London's theory, the wave functions of our large "atoms" would have to be of the order of magnitude of 137 times an ordinary interatomic distance. And yet we should have to have as many electrons per atom as usual, so that these electronic functions would have to overlap a great deal. This is not inconsistent, as a simple model will show.

Let us suppose that the electrons were entirely free, but that the box containing them was divided into a set of cubic cells by impenetrable walls. We imagine these cells to have sides of the order of 137 atomic diameters, and assume that there are enough electrons per unit volume so that there are of the order of $(137)^3$ electrons in each cell. Then all the wave functions will decrease to zero at the edge of the cell, so that they will fall off in such a way as we have described. The energy levels of electrons in such a cell would be similar to those given by the ordinary theory of free electrons in a metal, and the lowest $(137)^3$ of these levels would be occupied by the electrons, each extending throughout the cell. Thus we should have the same number of electrons per unit volume as before, but with wave functions falling off in a distance of the desired order, many electrons occupying the same region. With the radius 137 times that of an atom, the diamagnetism could be of the order of $(137)^2$ times as great, which is roughly what is required to produce the complete diamagnetism of London's theory. The energy levels, furthermore, are now discrete. The total width of the occupied levels will be very closely the same as in the absence of the barriers, or of the order of a few volts. The energy difference between levels will then be of the order of a few volts divided by $(137)^3$, or the order of 10^{-6} volts, a very small energy difference, of the order of magnitude of kT for 0.01° Abs. Hence we may expect the temperatures required to produce changes in this state to be correspondingly low. It thus seems possible by a simple model to secure something like London's theory, and this model has essential points in common with that to which our theory of the metal has led us.

A difficulty will at once occur to the reader. It

is not plain from this argument why we could not let the size of our cells increase, from 137 atom diameters, to any desired size. Then the diamagnetic susceptibility should increase without limit, as $[r^2]_{Av}$ continues to increase. And yet the model would approach the ordinary case of free electrons, from which we know by the work of Landau[7] that the diamagnetism has a very small value. We shall now consider this paradox, and show that our model is in fact consistent with the usual theory, and that there are points in connection with this theory which have not previously been brought out. What we shall do essentially is to consider the problem of N free electrons in a box, in which we vary the size of the box from a very small size, of atomic dimensions, through our size of the order of 137 atomic dimensions, to the whole size of an ordinary crystal, keeping the number of free electrons per unit volume (rather than the total number of free electrons) approximately constant. Then the energy separation of the levels will be of the order of a few volts divided by N, giving a few volts for a box as small as an atom ($N=1$), something like 10^{-6} for the box of the order of 137 atoms, as we have seen, and something like 10^{-23} volts for a whole crystal.

We shall consider first the two limiting cases of the single atoms and the entire crystal. The first is essentially the case of isolated atoms, or of the diamagnetism of bound electrons as treated by Peierls.[8] The second is the case of free electrons as treated by Landau.[7] In both cases the diamagnetism is small, and the methods used for treating them are similar in some respects. Hence it has been assumed naturally that one could interpolate between them, and that there was no place for large diamagnetism in the theory. We shall find that this view is erroneous, that the treatments of the two cases are essentially different, and that we make no error in supposing a large diamagnetism in our case of $N=(137)^3$.

Assume that a magnetic field is present, given by $\mathbf{H}=\text{curl }\mathbf{A}$, where $\mathbf{A}$ is a divergenceless vector potential. Then the relation between velocity $\mathbf{v}$ and momentum $\mathbf{p}$ is

$$m\mathbf{v}=\mathbf{p}-(e/c)\mathbf{A}=-(h/2\pi i)\nabla-(e/c)\mathbf{A}, \quad (2)$$

the second equation representing the operator form. In this case Schrödinger's equation $(\frac{1}{2}mv^2+V)\psi=E\psi$ becomes

$$\left\{-\frac{h^2}{8\pi^2 m}\nabla^2-\frac{ieh}{2\pi mc}\mathbf{A}\cdot\nabla+\frac{e^2A^2}{2mc^2}+V-E\right\}\psi=0. \quad (3)$$

For a constant magnetic field of magnitude H in the z direction, we can set $A_x=-Hy/2$, $A_y=Hx/2$, $A_z=0$, giving lines of $\mathbf{A}$ circulating around the z axis. Then Eq. (3) becomes

$$\left\{-\frac{h^2}{8\pi^2 m}\nabla^2-\frac{ieh}{4\pi mc}H\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)+\frac{e^2}{8mc^2}H^2(x^2+y^2)+V-E\right\}\psi=0. \quad (4)$$

For a wave function possessing no permanent magnetic moment, the second term, proportional to H, vanishes, so that since we are not considering paramagnetism we can leave it out of account. The effect of the magnetic field, then, is to add the term in H^2. This has quite different results in the two limiting cases we are considering.

For a single atom, or a box so small as to contain only one atom, the term $(e^2/8mc^2)H^2(x^2+y^2)$ is negligible for any magnetic field at our disposal. It is convenient to express in terms of atomic units. We measure distances in terms of the radius a_0 of the first Bohr orbit, energies in terms of the Rydberg energy. The natural unit of H is that magnetic field for which $(eh/2mc)H$, or the energy of a Bohr magneton in that magnetic field, equals the Rydberg energy. This of course is an enormous magnetic field, and turns out to be approximately 3.8×10^8 gauss. In these units, the term is $\frac{1}{4}H^2(x^2+y^2)$. Then if x^2+y^2 is of the order of unity, and H of the order of 40,000 gauss, or 10^{-4} atomic units, this term in the energy is of the order of 10^{-8} atomic units, entirely negligible compared with the energy of an atom. Thus the term can be neglected in the wave equation, and we can determine the wave function by the original wave equation, so that to the first order there is no perturbation in the wave function produced by the magnetic field. This is the analog in wave mechanics to the Larmor theorem in classical mechanics. The change in energy in the magnetic field is then to be found from the

[7] L. Landau, Zeits. f. Physik **64**, 629 (1930).
[8] R. Peierls, Zeits. f. Physik **80**, 763; **81**, 186 (1933).

average of the perturbation energy $\frac{1}{4}H^2(x^2+y^2)$ over the unperturbed orbit, and is thus proportional to H^2 (as is proper for diamagnetism) and to the mean square radius of the orbit. This is the basis of the ordinary theory of diamagnetism of atoms.

The situation is reversed in the case of free electrons in a box of the order of magnitude of the size of real crystals. In this case, the maximum x in the box will be of the order of 10^8 atomic units, so that if H is of the order of 40,000 gauss the magnetic energy will be of the order of 10^8 atomic units at the edge, an enormous value. In this case, in other words, the magnetic field furnishes a potential in which the electron must move, and the wave functions to a first approximation must be determined in this potential, and are different from those of free electrons. As is well known in the Landau solution, the corresponding wave functions resemble those of linear oscillators surrounding $x=y=0$ in the xy plane, falling off exponentially as the coordinates depart from these values, but are periodic with respect to z. This corresponds physically to the classical motion of the electrons, in which the z component is unaffected by the field, but the electrons spiral around the z axis, executing a periodic motion in the xy plane, which must be quantized in the quantum theory. For a given value of the z component of momentum, this quantization leads to an energy $2H(n+\frac{1}{2})$, where n is an integer, associated with the xy motion, in these units. In this case, the wave functions are changed in the first order by the presence of the magnetic field. It is then a complicated process to find the change in the energy of the whole system with the magnetic field. As the field changes, electrons have transitions from one energy level to another, and when the whole calculation is made, Landau's formula for diamagnetism of free electrons is found.

It is clear that the theories in these two cases are very different in nature. Now let us ask, what happens in the intermediate cases? If we let our box increase in size from atomic size, at what point must we begin to consider the first order effect of the field on the wave function? Up to this point we are allowed to use the atomic formula for diamagnetism, and the diamagnetism per electron will increase proportionally to the square of the dimension of the box. We shall not set up and solve this problem, though it should not be impossibly hard to handle it. But we can easily find the qualitative criterion to apply. It certainly seems most probable that if the spacing of levels as determined by the magnetic term in the potential energy (namely $2H$, in atomic units), is small compared to the spacing of the unperturbed levels (of the order of $1/N$, in atomic units), then we may use the unperturbed wave functions, and we have the atomic formula for diamagnetism. On the other hand, if $2H$ is large compared with $1/N$, we must find the new wave functions, treating the magnetic field as the principal term in the potential, and we shall get something much more like the free electron theory.

It is instructive to consider the variation, not with N at fixed H, but with H at fixed N. For any size box, and small enough magnetic field, the whole box will act like a big diamagnetic atom, but as the magnetic field increases a point will be reached where this is no longer true. Without a complete solution we cannot say exactly how the energy will act as function of H. In the region where the atomic formula holds, the energy will increase quadratically with H. When the formula breaks down it will presumably still increase, but much more slowly, so that the slope of the curve of energy against H, which gives the magnetic moment, will not only not increase proportionally with H, but may perhaps even decrease. As much larger fields are reached, the curve will approach another parabolic curve, with much smaller coefficient, corresponding to free electron diamagnetism. Physically, it is clear what is happening. At low fields, the Larmor precession is slow enough so that its frequency is negligible compared with the frequency of oscillation of an electron from one side of the box to another, its energy small compared to the energy difference between electrons. In particular, the term $\frac{1}{4}H^2(x^2+y^2)$ represents the increased kinetic energy on account of the Larmor precession, and this is negligible. At high fields, however, this added kinetic energy becomes large compared with the energy already present, and the whole motion is profoundly affected by the magnetic field. It is affected in such a way that the energy does not increase nearly as fast as the atomic

picture would indicate. This occurs by having the various electrons move in orbits much smaller than the box as a whole, either in the spirals of free electrons, or surrounding the atomic nuclei in a real crystal, each orbit executing its own Larmor precession about its center. As a result, no part of the electron's path acquires a great enough velocity from the precession to add greatly to the kinetic energy.

The question now becomes, if N is of the order of $(137)^3$, what is the order of magnitude of the magnetic field necessary to destroy the highly diamagnetic state? In this case, the spacing of levels is of the order of magnitude of $1/N=4\times10^{-7}$ atomic units. To get the maximum value of x^2+y^2, we may assume the interatomic distance to be about five atomic units (about 2.5A), so that with a radius of the cell of the order of 137 (5)/2, we could have x^2+y^2 up to about $(137)^2\ 25/4=1.2\times10^5$. Then if we let $\frac{1}{4}H^2(x^2+y^2)=1/N$, we have $H=3.6\times10^{-6}$ atomic units $=1400$ gauss. This very crude calculation, in other words, has given us a critical magnetic field of the order of magnitude of the fields actually observed to destroy superconductivity; it is obvious that in such a rough calculation a factor of ten can be disregarded.

We are finally led to this very significant result, then. If we assume large wave functions for electrons in a superconducting body, which become occupied at low temperature in the superconducting state, then the theory of diamagnetism as applied to them will predict that they should be perfectly diamagnetic, which London's theory has shown to be essentially equivalent to superconductivity. The theory predicts, however, that this should hold only for low magnetic fields. And the order of magnitude of the limiting field calculated in this way agrees in a rough way with that actually found to destroy superconductivity.

Of course, it is obvious that the electrons will not all really have orbits of the same size. Some will be smaller than the size necessary to produce perfect diamagnetism, some bigger. We cannot expect the theory to lead to orbits just the right size, except by coincidence. But it will do no serious harm if they are somewhat too big, for diamagnetism can clearly not be carried beyond the point where the magnetic field within the crystal is completely cancelled by the induced currents. If the orbits were somewhat bigger than we have calculated, on the average, the effect would be to produce the transition at a lower critical magnetic field, more nearly the experimental situation. Also it is to be noted that the type of theory we have spoken of here would not lead to a sudden change of properties with increasing field, but instead to a gradual transition. We cannot expect to understand the sudden change except by combining these magnetic considerations with a real long-scale cooperative theory, for it is only in this way that sharp transitions can be explained. This would have to be done in a more elaborate theory.

Finally, to make the discussion more intelligible, we should point out, as London has done, the way in which the current appears in this form of theory. Eq. (2) for the velocity has a term $-(e/mc)A$, which is not present when the magnetic field is absent. This term appears in the presence of the field without any change in the wave function. Thus in our case of low fields, where the wave function is unchanged, there will be a term in the current, proportional to the magnetic field. In a uniform field, A circulates about the axis of the field, so that this additional current will do the same. The current thus produced is $-(e^2/mc)\psi^*\psi A$, a constant times A, which is London's fundamental equation. It is then clear that the resulting current produces a magnetic moment of the sample as a whole. More complicated cases will not be discussed in detail, for London has fully treated the consequences of this type of current. As London has pointed out, a current produced in this way by the magnetic field is not subject to dissipation by interaction with the lattice, so that this conduction is of essentially a different nature from ordinary conduction, and is not subject to resistance of the ordinary sort at all.

The writer is indebted to Professor H. Grayson Smith for helpful correspondence, and to Professors E. Teller and J. H. Van Vleck for valuable conversations on the subject of this paper. He is also greatly indebted to the Institute for Advanced Study for the privilege of conducting these investigations in residence at Princeton.

The Intensity and Polarization of the Light Diffracted by Supersonic Waves in Solids

HANS MUELLER
George Eastman Research Laboratory of Physics,
Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received May 7, 1937)

The intensity and polarization of the light diffracted by supersonic waves in solids are calculated by considering the photoelastic effect due to the strains created by the elastic waves. For natural incident light the diffracted light is partially polarized. In glasses and cubic crystals the diffracted light is partially depolarized if the incident light is polarized. The intensities are evaluated with the help of the theory of Raman and Nath and it is shown that this theory explains the measurements better than Brillouin's approximation. The calculated intensity distributions for natural and polarized incident light agree with the observations of Schaefer and Bergmann on glasses, quartz and calcite. This agreement justifies Fues' and Ludloff's assumption that all forced vibrations of the same frequency oscillate with the same amplitude.

SCHAEFER and Bergmann[1] have shown that supersonic waves in crystals act as diffraction gratings for light. Their results justify the assumption that the piezoelectric oscillator excites over a narrow frequency range all the standing waves which can exist in a crystal according to the theory of specific heats. Fues and Ludloff[2] have given a theory of the diffraction pattern. They calculate the wave-lengths of the infinite number of progressing waves of a given frequency which can travel in an infinite crystal, and then consider the diffracted light as due to Bragg reflections on the wave fronts. Since the positions of the interference maxima are the same for standing and traveling waves this theory gives the locus of all diffraction spots produced by the finite number of standing waves in a finite crystal.

In some diffraction patterns all spots appear to be of nearly equal intensity. This observation led Fues and Ludloff to conclude that the coupling between the different modes of vibration of a crystal results in an equipartition of energy for all sound waves of the same frequency. Since this conclusion, if correct, is of considerable theoretical importance, and since no theoretical proof of its validity has been given, it seems worth while to see whether it really is justified by the experimental evidence. Many of the published photographs of the diffraction patterns, particularly those taken with polarized light, reveal a systematic variation of the light intensities along the diffraction curves. Some parts of the calculated curves are not observed. A test of Fues and Ludloff's hypothesis calls therefore for an investigation of the intensities of the interference maxima. We will show below that a large part[3] of the observed intensity variations can be accounted for by assuming that all waves oscillate with approximately the same amplitude. Hence the proposed principle of equipartition appears to be validated by the observations.

The intensity of the diffracted light depends primarily on the periodic variations of the refractive index which are produced by the sound waves. The calculation of these index variations is a problem of photoelasticity. An elastic wave creates periodically varying strains which in turn introduce local periodic alterations of the optical index ellipsoid. These pulsations of the index ellipsoid determine the fluctuations of the index of refraction for light polarized in any direction. As was first shown by Brillouin[4] the angle of deviation of the diffracted light is given by Bragg's law, but the intensities must be calculated by using the theories developed by Raman and Nath[5] and Estermann and

[1] C. Schaefer and L. Bergmann, Naturwiss. **22**, 685 (1934).

[2] E. Fues and H. Ludloff, Sitzungsber. d. preuss. Akad., Math. phys. Kl. **14**, 222 (1935).

[3] Some diffraction patterns, for instance those for glasses (C. Schaefer and L. Bergmann, Naturwiss. **23**, 799 (1935)), show also intensity variations which apparently are due to the method of excitation of the sound waves. The waves which are directly transmitted from the piezo quartz to the glass or crystal are much stronger than all others.

[4] L. Brillouin, "La diffraction de la Lumière par des Ultra-sons," Act. Sci. et Ind. (1933).

[5] C. V. Raman and N. S. N. Nath, Proc. Ind. Acad. Sci. **2**, 406 (1935).

Wannier[6] for the Debye-Sears effect in liquids. Since these theories give the intensities as a function of the amplitude of the index variations and since their results are applicable to solids without any modifications we have to discuss here only the photoelastic problem.

The first part of the calculation involves the determination of the wave-lengths and directions of oscillation of all sound waves of a given frequency which can travel in an infinite crystal. Fues and Ludloff have given the solution of this problem. The three differential equations of a sound wave in a crystal

$$\rho\frac{\partial^2 u_x}{\partial t^2}=\sum_y\sum_{x'}\sum_{y'}c_{xx',\,y'y}\frac{\partial^2 u_y}{\partial x'\partial y'} \tag{1}$$

have the solutions

$$u=a\exp i[\omega t-(kr)] \tag{2}$$

provided the components of the amplitude vector satisfy the three linear homogeneous equations

$$\begin{aligned} &a_x(k^2Q_{xx}-\rho\omega^2)+a_yk^2Q_{xy}+a_zk^2Q_{xz}=0,\\ &a_xk^2Q_{yx}+a_y(k^2Q_{yy}-\rho\omega^2)+a_zk^2Q_{yz}=0,\\ &a_xk^2Q_{zx}+a_yk^2Q_{yz}+a_z(k^2Q_{zz}-\rho\omega^2)=0. \end{aligned} \tag{3}$$

The wave-length λ_s of the sound wave is determined by the condition that the determinant of Eqs. (3) must vanish.

$$|k^2Q_{xy}-\rho\omega^2\delta_{xy}|=0. \tag{4}$$

The wave vector $\mathbf{k}$ has the length $2\pi/\lambda_s$ and points in the direction of propagation. The Q_{xy}'s are quadratic forms

$$Q_{xy}=Q_{yx}=\sum_{x'}\sum_{y'}c_{xx',\,y'y}\alpha_{x'}\alpha_{y'} \tag{5}$$

where α_x, α_y, α_z are the direction cosines of $\mathbf{k}$, and $c_{xx',\,y'y}=c_{ik}$ are the elastic constants of the crystal.[7] To every direction of propagation there exist three solutions of Eq. (4). In an isotropic solid these correspond to one longitudinal and two transversal waves. In crystals the waves are usually neither longitudinal nor transversal, but the amplitude vectors of the three waves traveling in the same direction are always mutually perpendicular.

The solutions $\mathbf{k}(\alpha_x, \alpha_y, \alpha_z)$ of Eq. (4) determine the so-called "form-frequency surface." Fues and Ludloff have shown that the diffraction pattern is a curve geometrically similar to the intersection of this surface with a plane normal to the direction of the incident light. This means that only those waves which travel in directions normal to the light beam are effective in the diffraction.

By introducing the solutions of Eq. (4) in Eqs. (3) one finds the direction of oscillation of each wave. The absolute value of the amplitude remains of course undetermined. In accordance with the proposed principle of equipartition of energy we assume that the amplitude has the same value for all directions of oscillation and propagation.

After all possible sound waves in a crystal have thus been found, the next step is to calculate their photoelastic effects. Differentiation of Eq. (2) gives the strains due to the sound waves

$$x_x=\frac{\partial u_x}{\partial x},\quad z_y=y_z=\frac{\partial u_y}{\partial z}+\frac{\partial u_z}{\partial y}. \tag{6}$$

Let us assume that the axes x, y, z have been chosen to coincide with the axes of the crystal's optical index ellipsoid:

$$x^2/n_1^2+y^2/n_2^2+z^2/n_3^2=1,$$

where n_1, n_2, n_3 are the principal indices of refraction. The local strains will rotate and deform this ellipsoid and its equation takes the general form

$$\begin{aligned} &a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{23}yz\\ &\qquad+2a_{31}zx+2a_{12}xy=1. \end{aligned} \tag{7}$$

The polarization constants a_{ik} in the deformed state differ from those in the undeformed state by quantities which are to a first approximation linear functions of the local strains. Hence

[6] R. Estermann and G. Wannier, Helv. Phys. Acta **9**, 337, 520 (1936).

[7] Using Voigt's notation for the elastic constants the coefficients of the Q_{xy}'s are given by the following array:

	α_x^2	α_{y^2}	α_z^2	$\alpha_y\alpha_z$	$\alpha_z\alpha_x$	$\alpha_x\alpha_y$
Q_{xx}	c_{11}	c_{66}	c_{55}	$2c_{56}$	$2c_{15}$	$2c_{16}$
Q_{yy}	c_{66}	c_{22}	c_{44}	$2c_{24}$	$2c_{46}$	$2c_{26}$
Q_{zz}	c_{55}	c_{44}	c_{33}	$2c_{34}$	$2c_{35}$	$2c_{45}$
Q_{yz}	c_{56}	c_{24}	c_{34}	$(c_{23}+c_{44})$	$(c_{45}+c_{36})$	$(c_{25}+c_{46})$
Q_{zx}	c_{15}	c_{46}	c_{35}	$(c_{26}+c_{45})$	$(c_{13}+c_{55})$	$(c_{56}+c_{14})$
Q_{xy}	c_{16}	c_{26}	c_{45}	$(c_{25}+c_{46})$	$(c_{14}+c_{56})$	$(c_{12}+c_{66})$

$$a_{11}-1/n_1^2=p_{11}x_x+p_{12}y_y+p_{13}z_z+p_{14}z_y+p_{15}x_z+p_{16}y_x,$$
$$a_{23}=p_{41}x_x+p_{42}y_y+p_{43}z_z+p_{44}z_y+p_{45}x_z+p_{46}y_x. \quad (8)$$

The 36 constants $p_{ik}\neq p_{ki}$ are Pockels' elasto-optical constants.[8] They are known for a large number of glasses, where their number is reduced to two, for the cubic crystals NaCl, KCl and CaF_2 (3 constants) and for quartz and calcite (8 constants).

The first three equations of (8) determine the fluctuations of the lengths of the axes of the index ellipsoid, the last three give the change of orientation of these axes. The variations of the indices of refraction of the two plane polarized components of a light beam are determined, of course, by the variations in lengths of the axes of the ellipse which is formed by the intersection of the index ellipsoid with a plane normal to the direction of propagation of the beam.

For a triclinic crystal these calculations would be very laborious. They are, however, fairly simple for crystals of the rhombic, rhombohedric, hexagonal and cubic systems provided the light is incident parallel to one of the crystallographic axes. If, for instance, the light enters in the x direction and oscillates in the y direction the amplitude of the variation of the index of refraction is

$$\Delta n=\tfrac{1}{2}n_2^3(a_{22}-1/n_2^2), \quad (9)$$

where $(a_{22}-1/n_2^2)$ is obtained by introducing in Eq. (8) the amplitudes of the strains given by Eq. (6). This value of Δn determines the intensity of the diffracted light component which has the same polarization as the incident light. Due to the periodic change of the orientation of the index ellipsoid the diffracted beam will also have a component oscillating in the z direction. If the crystal is naturally birefringent for light traveling in the given x direction this "depolarizing" component along z is very small and can be neglected.

The situation is, however, entirely different if the light travels in the direction of an optical axis or in any arbitrary direction through a cubic crystal or an amorphous solid.[9] In this case the intersect of a plane normal to the light beam with the index ellipsoid is a circle for the undeformed state. The deformations change this circle to an ellipse whose axes have a *fixed* direction. For instance for light traveling in the x direction the angle θ between one axis of the ellipse and the y axis is given by

$$\tan 2\theta=2a_{23}/(a_{22}-a_{33}). \quad (10)$$

Hence one resolves the incident light into two components oscillating parallel to these axes. The amplitudes of their index variations are

$$\Delta n=\tfrac{1}{4}n^3[(a_{22}-1/n^2)+(a_{33}-1/n^2)\pm((a_{22}-a_{33})^2+4a_{23}^2)^{\frac{1}{2}}]. \quad (11)$$

Consequently for plane polarized incident light the diffracted beam is usually partly depolarized, and for natural incident light the diffracted beam is partially polarized.

As mentioned previously the relation between Δn and the intensity of the diffracted light has been investigated by Raman and Nath and Estermann and Wannier. If Δn is large the relation is very complicated and higher order diffraction images appear. If Δn is very small only the first order is visible and its intensity is proportional to $(\Delta n)^2$. In the experiments of Schaefer and Bergmann only the first order is observed, but this fact is not sufficient justification for the use of Brillouin's approximation $I=c(\Delta n)^2$. According to Raman and Nath[10] the best approximation for the intensity of the first order diffraction images due to standing supersonic waves is

$$I(v)=c\int_0^{2\pi}J_1^2(v\sin\psi)d\psi, \quad (12)$$

where J_1 is a Bessel function, $v=2\pi\Delta nL/\lambda$ and L/λ is the length of the optical path in the

[8] F. Pockels, *Lehrbuch der Kristalloptik* (Leipzig, 1906) p. 460. For amorphous solids the p_{ik} can be replaced by Neumann's strain-optical constants $q=\frac{1}{2}np_{11}$, $p=\frac{1}{2}np_{12}$, $p_{44}=\frac{1}{2}(p_{11}-p_{12})$.

[9] The intensity ratio of the "depolarizing" and polarized components is proportional to $a_{23}^2/(a_{22}-a_{33})^2$. For birefringent crystals the denominator differs very little from $(1/n_2^2-1/n_3^2)^2$. Hence the ratio is proportional to the square of the small strains x_x, y_z etc. For cubic crystals, however, where $n_1=n_2=n_3$ the denominator is proportional to the square of the strains and hence the above ratio is of the order of magnitude 1.

[10] C. V. Raman and N. S. N. Nath, Proc. Ind. Acad. Sci. **3**, 75 (1936).

crystal. For $v<0.2$ this relation is identical with Brillouin's approximation. For $v>2$ the higher orders have an appreciable intensity. Since they do not appear in the experiments we are only interested in the range where $v<2$. In this range $I(v)$ increases continuously (a maximum is reached at about $v=2.2$). Hence we conclude that the observed intensities must be a monotonic function of Δn and that the intensity variations are analogous to the variations of Δn.

At present only a qualitative verification of the theory is possible because no intensity measurements are available. We will show below that all intensity variations observed in the diffraction patterns of glasses, cubic crystals, quartz and calcite agree qualitatively with the variations of $i=2\Delta n/n^3$. The variation of this quantity can be calculated with the help of Eq. (9) or (11) from the elastic and elasto-optical constants of the solid. The absolute value of i is proportional to $a\omega\sqrt{\rho}$, where a is the amplitude of the sound wave, $\omega=2\pi\nu$ its frequency and ρ the density of the solid. The data given below are calculated[11] with $a\omega\sqrt{\rho}=10^7$.

Isotropic Solids

For isotropic solids the diffraction pattern consists of two concentric circles. The ratio of their radii is $k_1/k_2=[\frac{1}{2}(1-2\sigma)/(1-\sigma)]^{\frac{1}{2}}$, where σ is Poisson's ratio. Let x be the direction of the incident beam, y the direction of oscillation of the plane polarized light, and consider a diffraction spot on the diameter which forms an angle φ with the y axis.

The inner circle is due to the longitudinal waves. The Eqs. (8), (10) and (11) furnish $\theta=\varphi$, $i_1=p_{11}=2q/n$, $i_2=p_{12}=2p/n$, where i_1 refers to the light component which oscillates in the direction θ and i_2 to the component normal thereto.

The outer circle is a double solution since it corresponds to the two transversal sound waves. For those transversal waves which oscillate in the x direction we get $i_1=i_2=0$. Hence they do not contribute to the diffraction pattern. The outer circle is produced exclusively by those shear waves which oscillate in the yz plane.[12] For these waves $\theta=\varphi+\pi/4$, $i_1=i_2=p_{44}=(q-p)/n$.

It is now a simple matter to find the intensity distributions for any experimental condition. If we denote with A, B, C the values of the intensity function (12) for the arguments $v_A=4\pi n^2k_1aqL/\lambda$, $v_B=4\pi n^2k_1apL/\lambda$ and $v_C=2\pi n^2k_2a(q-p)L/\lambda$, where k_1 and k_2 are the lengths of the wave vectors for the longitudinal and transversal waves and p and q are Neumann's strain-optical constants, we have:

For observations with crossed Nicols

Inner ring $I=\frac{1}{4}(A+B)\sin^2 2\varphi$,
Outer ring $I=\frac{1}{2}C\cos^2 2\varphi$.

For observations with parallel Nicols

Inner ring $I=A\cos^4\varphi+B\sin^4\varphi$,
Outer ring $I=\frac{1}{2}C(1+\sin^2 2\varphi)$.

For observations with a single Nicol the results are the same whether it is inserted in the incident or transmitted beam

Inner ring $I=A\cos^2\varphi+B\sin^2\varphi$,
Outer ring $I=C$.

For observations with natural light all spots on each ring have of course the same intensity. The intensity ratio between spots on the inner and outer circle is $(A+B)/2C$.

The experimental results on glasses are in good agreement with the calculations. For glasses $(q-p)$ is always much smaller than either q or p and it may even vanish. Hence the outer circle appears always much weaker than the inner one.[13] The photographs made with crossed Nicols[14] show clearly that the intensity vanishes for $\varphi=0°$ and $90°$ on the inner circle, and for $\varphi=\pm 45°$ on the outer circle, in agreement with the above equations. For parallel Nicols the intensity variations are less pronounced but they follow the theoretical predictions.

[11] In the experiments ($\rho\omega^2=10^{16}$, $a^2=10^{-14}$) $a\omega\sqrt{\rho}$ is probably about 10, hence i is of the order of magnitude 10^{-5} (see Table I and the figures). For a crystal 1 cm thick $v=\pi n^3 i/\lambda$ is therefore approximately 1 and our assumption $v<2$ appears tenable.

[12] The discovery o diffraction patterns produced by transversal waves is due to E. Hiedemann and K. H. Hoesch, Zeits. f. Physik **98**, 141 (1935). The statement of Fues and Ludloff that transversal waves should not produce a diffraction pattern, since they create no density fluctuations, is of course incorrect. In solids a shear creates index variations without changing the density. On the other hand the diffraction patterns from the longitudinal waves are not solely due to density variations. The values of p and q depend also on other effects, as was shown by the author, Physics **6**, 179 (1935).

[13] $(q-p)$ has the largest values for quartz glass and light crown glasses and is very small for the heaviest flints. All published photographs were made with light glasses. We suspect that it is very difficult to observe the outer ring if heavy glasses are used.

[14] C. Schaefer and L. Bergmann, Naturwiss. **23**, 799 (1935).

Recently Bergmann[15] has measured the ratio A/B for a series of different glasses. His data show definitely that A/B is not proportional to $(q/p)^2$, as Fues[15] assumes. Fues' result would be correct if Brillouin's approximation were applicable. The fact that it is inadequate shows that the Raman-Nath theory is of importance not only for liquids but also for solids. The value of A/B depends on the amplitude of the soundwave and if p and q were known Bergmann's data could be used to determine this amplitude.

Cubic Crystals

The calculations have been carried out for NaCl, KCl and CaF_2 using the elastic constants of Voigt and the elasto-optical constants of Pockels.[16] If the light travels in a direction x normal to a cube face the observed diffraction patterns consist of two rings, which are not circular but have the symmetry 4 mm (C_{4v}). A third ring, a circle, which should be present according to the theory of Fues and Ludloff, does not appear. It can easily be shown that the intensity of this circle is zero, because it would be due to transversal waves vibrating in the x direction. The strains created by such a wave do not alter the cross section of the index ellipsoid normal to the x axis and hence $i_1=i_2=0$. The vibrations which give rise to the inner ring are "quasi-longitudinal," i.e., their direction of oscillation forms a small angle with the direction of propagation. Similarly the outer ring is due to "quasi-transversal" waves oscillating in the yz plane. In Table I we give the quantities θ, i_1, i_2, defined in the previous sections, as functions of the angle φ which determines the position of the diffractions spots on the rings. Due to symmetry conditions φ has to be varied only between 0 and 45°. The intensities of the diffracted light are given by the formulas: $I(\varphi)=\frac{1}{4}(A+B)\sin^2 2\theta$ for crossed Nicols, $I(\varphi)=A\cos^4\theta+B\sin^4\theta$ for parallel Nicols, the Nicols being parallel to the crystal axes. For natural incident light the diffracted light is partially polarized and has the intensity $(A+B)$. A and B are the values of the intensity function (12) for the arguments

$$v_A=\pi n^3 i_1 L/\lambda, \qquad v_B=\pi n^3 i_2 L/\lambda.$$

Table I. *The orientation and amplitudes of vibration of the index ellipse due to supersonic waves in cubic crystals.*

	NaCl			KCl			CaF_2		
φ	θ	i_1	i_2	θ	i_1	i_2	θ	i_1	i_2
Inner ring									
0	0	62.8	81.6	90	86.0	122.0	90	55.8	13.5
13.4	6	63.5	81.0	81	87.7	118.1	87	55.3	15.4
19.6	10	65.2	81.6	76	89.7	118.8	86	54.2	17.4
25.4	15	66.5	81.5	71	91.8	121.2	83	52.7	20.1
31.7	22	68.4	81.0	64	96.3	126.0	79	50.0	24.0
45	45	70.8	81.2	45	100.6	133.4	45	44.6	30.8
Outer ring									
0	45	9.5	9.5	45	34.4	34.4	45	12.7	12.7
13.4	65	17.4	4.5	35	1.2	61.2	17	13.4	26.1
19.6	72	19.9	4.7	30	9.0	66.7	12	16.5	32.1
25.4	77	21.6	5.3	25	13.5	67.0	8	20.2	36.1
31.7	84	20.4	8.6	18	13.3	63.3	5	24.7	38.5
45	90	15.6	15.6	0	23.4	23.4	0	35.2	35.2

Bergmann[17] has published photographs of the diffraction patterns for all three crystals. They were made with natural light and show very little intensity variations. This agrees with the results in Table I because wherever i_1 is small i_2 is large and *vice versa*, and hence the sum of the intensities of both components varies but little. The observations also agree with the theory insofar as the outer ring appears less intense than the inner one. No observations with polarized light have been published. They should show pronounced intensity variations. Owing to the fact that NaCl, KCl and CaF_2 belong to three different photoelastic groups[18] these variations must have a different character for the three crystals. *Vice versa* it will be possible to decide to which photoelestic group other cubic crystals belong by studying the diffraction patterns with parallel and crossed Nicols.

Quartz and Calcite

Both crystals belong to the rhombohedric class. Hence the calculations involve 14 parameters

[15] L. Bergmann and E. Fues, Naturwiss. **24**, 492 (1936). Although the strain-optical constants of the glasses used by Bergmann are not known, their values can be estimated by using the formula of H. Mueller, Physics **6**, 179 (1935).

[16] Except for KCl where the elastic constants have been corrected to satisfy Cauchy's relation. The diffraction pattern of KCl offers a very simple method for the verification of this relation.

[17] L. Bergmann, Physik. Zeits. **37**, 867 (1936).

[18] Hans Mueller, Phys. Rev. **47**, 947 (1935).

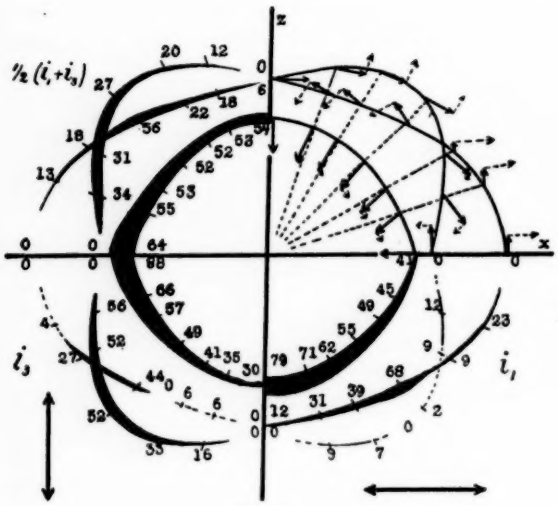


FIG. 1.

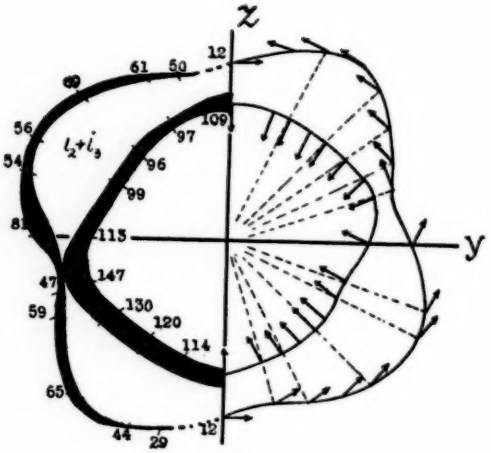


FIG. 2.

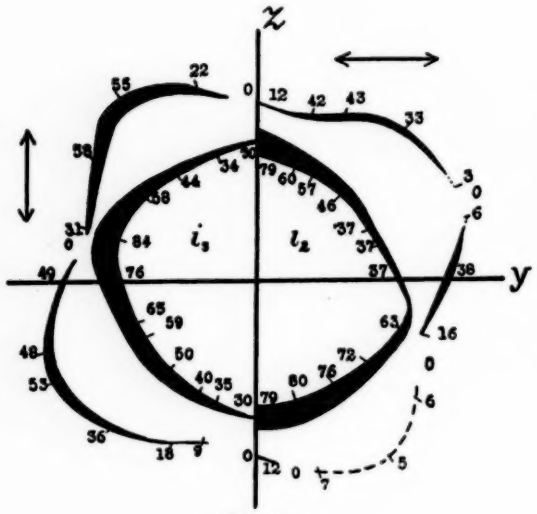


FIG. 3.

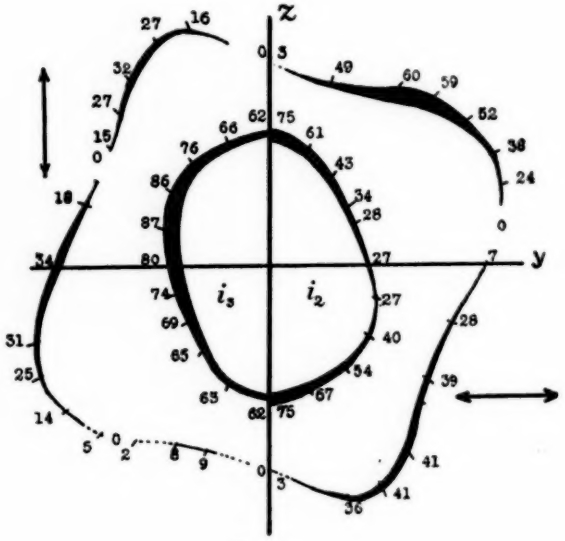


FIG. 4.

FIG. 1. Quartz, light passing in y direction. The arrows in the first quadrant give the directions of oscillation of the sound waves traveling normal to the y axis. The dotted vectors are the components of the amplitude vectors in the y direction. Second quadrant: the intensity distribution for natural incident light. Third quadrant: for polarized light oscillating in z direction. Fourth quadrant: for polarized light oscillating in x direction.

FIG. 2. Quartz, light passing in x direction. Right side: directions of oscillation of the sound waves traveling normal to x axis. Left side: intensities for natural incident light.

FIG. 3. Quartz, light traveling in $-x$ direction. Right side: intensities for light polarized in z direction. Left side: for light polarized in y direction.

FIG. 4. Calcite, light passing in x direction. Right side: intensities for polarized light oscillating in y direction. Left side: for light oscillating in z direction.

(6 elastic constants and 8 elasto-optical constants). Since the calculations, as outlined, are based on well-known principles of classical crystal physics they are of little interest and can be omitted. We give therefore only the results for those cases where a comparison with published data is possible.

For light passing through quartz in the direction of the optical axis the observed intensity distribution is practically uniform and

is not changed appreciably by using polarized light. While the theory predicts a noticeable variation for polarized light, it does not take into account the optical activity of quartz, and this smoothes out any intensity variation.

For light traveling in the y direction the diffraction pattern for a rhombohedric crystal is symmetrical with respect to the x and z axes. Hence it is sufficient to give the results for one quadrant only (Fig. 1). If the incident light is plane polarized in either the x or z direction the diffracted light has the same polarization. For natural incident light the diffracted light is partially polarized and its intensity is the sum of the intensities of the two components. The vectors in the first quadrant of Fig. 1 give the directions of oscillation of the supersonic waves. The inner ring is due to quasi-longitudinal waves, the outer rings are produced by quasi-transversal waves. The numbers in the second, third and fourth quadrants are the absolute values of $\frac{1}{2}(i_1+i_3)$, $i_3=a_{33}-1/n_3^2$ (Eq. (9)) and $i_1=a_{11}-1/n_1^2$, respectively. Guided by the form of the intensity function (12) we have sketched the expected intensities for natural and for polarized light.

Figures 2, 3 and 4 give the results for light traveling in the x direction through quartz and calcite. Since x is a polar axis the diffraction pattern changes to its mirror image if the light direction is reversed (Figs. 2 and 3). The patterns consist of two rings which have central symmetry. Hence in all figures only half the pattern is sketched for each case. The theory of Fues and Ludloff predicts the existence of a third ring, an ellipse, which is due to transversal waves oscillating in the x direction. This ellipse has everywhere the intensity zero.

When comparing these theoretical curves with the observations[1] one must take into account that in the experiments the patterns consist of a finite number of spots, corresponding to the finite number of standing waves in a finite crystal. It may therefore happen that no spots appear in places where the calculation predicts a large intensity. The results must be judged by the fact that strong spots appear only where the calculated intensity is large and that no spots are visible where the theory gives a very small intensity. In this respect the agreement is very satisfactory indeed. All the peculiarities of the complicated intensity distributions agree in every detail with the published photographs. We believe that this is sufficient proof that both the theory and the proposed principle of equipartition of energy are correct. A rigorous test, of course, involves intensity measurements.

The agreement between theory and observation may be considered a verification of Pockels' values of the elasto-optical constants. Bergmann and Fues[15] have pointed out that the intensities of the diffracted light might be used to determine these constants. This method is, however, only feasible after the intensity function (12) has been verified, or if the experiments are carried out with small enough sound amplitudes for which Brillouin's approximation is valid.

Since, as we have shown, the light diffracted by forced oscillations of solids is usually depolarized, it is natural to expect that also the light diffracted by the standing waves representing the temperature motion will be depolarized. This suggests a new explanation of the results of Krishnan[19] on the light scattering in glasses, which would not involve his assumption of crystalline groupings.

---

[19] R. S. Krishnan, Proc. Ind. Acad. Sci. **3**, 211 (1936). The variation of the depolarization $1/\rho_h$ with the chemical composition is similar to the behavior of the photoelastic constant $(q-p)$.

Internal Friction in Solids

I. Theory of Internal Friction in Reeds

CLARENCE ZENER
Wayman Crow Hall of Physics, Washington University, St. Louis, Missouri

(Received April 23, 1937)

In a vibrating reed opposite sides have dilations of opposite signs. Thus when one side is heated the other is cooled. At low frequencies the vibrations are isothermal. At high frequencies they are adiabatic. At intermediate frequencies they are of a hybrid type accompanied by internal friction. In this paper this internal friction is calculated solely from thermodynamical considerations. It is predicted that the internal friction associated with this hybrid type of vibration is of a larger order of magnitude than that due to all other causes.

§1. INTRODUCTION

IN spite of numerous experimental investigations of the internal friction in solids, no satisfactory theoretical study has been given. Our ignorance as to the mechanism of internal friction is due partly to the fact that no one mechanism is responsible for the energy losses in all cases of vibrating solids. As an example of the complexity of the phenomena, when the strains are increased beyond a critical value the losses increase more rapidly than as the square of the strains.[1]

In searching for possible mechanisms of internal friction in solids, it is natural to seek suggestions from the much studied field of internal friction in gases. Here sound waves are damped most strongly in bands whose periods of vibration are comparable to the relaxation times of the gas.[2] For example, sound waves in N_2 are damped most strongly for those periods of vibration which are comparable to the time required for temperature equilibrium to be established between the vibrational and the translational degrees of freedom of the N_2 molecules. We are thus led to investigate types of thermal nonequilibria within a solid.

The following possibilities present themselves. (1) A close analogy to the above mentioned case of N_2 gas would arise from the unequal cooling during expansion, or the unequal heating during compression, of the Debye normal modes of different frequencies. Thermal equilibrium between the different normal vibrations is, however, established so rapidly that damping would occur only for frequencies much higher than those at present available. (2) A lack of thermal equilibrium exists between the vibrating solid and the surrounding air. In an investigation on springs, Sayre[3] found that the partial lack of thermal equilibrium between a wire and the surrounding air gives rise to the dominant energy losses in certain cases. Since the time required for the establishment of temperature equilibrium between a wire and the air is comparatively long, such losses are confined to very low frequencies. (3) We come finally to the lack of thermal equilibrium between various parts of the vibrating solid itself. In isotropic solids with a positive thermal expansion coefficient, a positive (cubic) dilation lowers the temperature, a negative dilation raises the temperature. However, just as in a gas, the time required for the establishment of thermal equilibrium between regions separated by a half wave-length is much longer than the period of vibration. A solid differs from a gas in that fluctuations in dilation, and hence also in temperature, may exist between regions much closer together than a half wave-length. For example, the two sides of a reed vibrating transversely have dilations of opposite signs. Again, the dilation varies rapidly in the immediate vicinity of imperfections in the solid.

The purpose of this paper is to investigate whether the local fluctuations in temperature within a vibrating solid may be of importance in

[1] See Voight, Ann. d. Physik **47**, 671 (1892); Wegel and Walther, Physics **6**, 141 (1935).
[2] P. S. H. Henry, Proc. Camb. Phil. Soc. **28**, 249 (1932).
[3] Sayre, Rheology **3**, 206 (1932).

internal friction. After defining the various yardsticks of internal friction (§2), the effect of the temperature fluctuation across a vibrating reed upon its internal friction is calculated (§3). The results are discussed quantitatively in §4. It is predicted that over a wide frequency band the effect here discussed will be the predominant cause of internal friction. This thermodynamic effect disappears at low temperatures.

A direct experimental check of this theory may be obtained by observing whether the internal friction of a reed has a maximum in the frequency range predicted by Eq. (11).

§2. Yardsticks of Internal Friction

The displacement vector $\mathbf{U}(\mathbf{r}, t)$ in a vibrating solid is determined by a differential equation of the form

$$\partial^2\mathbf{U}/\partial t^2+L\mathbf{U}=\mathbf{F}+\mathbf{f}, \tag{1}$$

and by a set of boundary conditions. Here the applied body force (per unit mass) is represented by $\mathbf{F}$. The term $-L\mathbf{U}$ represents the force acting upon unit mass due to the elasticity of the solid. The elastic coefficients in the operator L are taken to be the isothermal coefficients. The forces, per unit mass, which arise from temperature changes,[4] from viscosity, etc., are denoted by $\mathbf{f}$. In particular, the surface forces due to the air are taken care of by a term in $\mathbf{f}$ which is zero everywhere except near the surface.

We shall examine first the solutions of (1) when F is a periodic function of time. We eliminate time from this equation by the usual artifice of complex symbols. Using German characters to denote vectors which are functions of the coordinates, but not of time, we set

$$\mathbf{F}(\mathbf{r}, t)=\mathfrak{F}(\mathbf{r})e^{i\omega t},$$
$$\mathbf{f}(\mathbf{r}, t)=\mathfrak{f}(\mathbf{r})e^{i\omega t},$$
$$\mathbf{U}(\mathbf{r}, t)=\mathfrak{U}(\mathbf{r})e^{i\omega t}.$$

Equation (1) now becomes

$$(L-\omega^2)\mathfrak{U}=\mathfrak{F}+\mathfrak{f}. \tag{2}$$

In general $\mathbf{U}$ and $\mathbf{f}$ are not in phase with each other. In fact, it is just this difference in phase between $\mathbf{U}$ and $\mathbf{f}$ which gives rise to damping. If we choose the origin of time so that $\mathfrak{U}$ is real, $\mathfrak{f}$ will be complex. Let $\mathfrak{f}_1$, $\mathfrak{f}_2$ be its real and imaginary parts, respectively, i.e.,

$$\mathfrak{f}=\mathfrak{f}_1+i\mathfrak{f}_2.$$

Further let $\mathfrak{U}_0$ be the solution of (2) when $\mathfrak{f}$ is neglected. Since the force $\mathfrak{f}$ has only a small effect upon the solution of (2), it may be calculated assuming the displacement of the solid to be given by $\mathfrak{U}_0$. Then the ratio

$$Q^{-1}=\frac{|\int\mathfrak{U}_0\cdot\mathfrak{f}_2dv|}{|\int\mathfrak{U}_0\cdot L\mathfrak{U}_0dv|} \tag{3}$$

is a measure of the internal friction of a solid, the integration being over the entire solid. The definition (3) is analogous to the definition of the Q of an electrical circuit. Here, as in the electrical analogy, Q is in general a function of frequency.

When ω is nearly equal to one and to only one natural angular frequency, ω_k, the *form* of the vector function $\mathfrak{U}$ will be nearly identical to that of a vibration vector $\mathfrak{U}_k$ of the free undamped oscillation. This vector satisfies the equation

$$(L-\omega_k^2)\mathfrak{U}_k=0.$$

The *amplitude* of $\mathfrak{U}$ will depend upon the impressed frequency $\omega/2\pi$. We shall take $\mathfrak{U}_k$ to be normalized, i.e.,

$$\int\mathfrak{U}_k\cdot\mathfrak{U}_kdv=1,$$

and set $$\mathfrak{U}=C\mathfrak{U}_k,$$

where C is a function of the impressed angular frequency ω. The force $\mathfrak{f}$ will then also be proportional to C. We shall denote by $\mathfrak{f}_k$ that force which is associated with the normalized function U_k. Then we may set

$$\mathfrak{f}=C\mathfrak{f}_k.$$

We substitute these expressions into (2), multiply to the left by $\mathfrak{U}_k$, and integrate over the solid. We obtain

$$C(\omega)=\frac{F_k}{(\omega_k^2-\omega^2)+i\eta_k},$$

where $$F_k=\int\mathfrak{U}_k\cdot\mathfrak{F}dv,$$

and $$\eta_k=\int\mathfrak{U}_k\cdot\mathfrak{f}_{k2}dv.$$

[4] Love, *The Mathematical Theory of Elasticity* (Cambridge Press, 1920), p. 106.

A plot of $|C_\omega|^2$ against ω has a sharp maximum at $\omega=\omega_k$. The width of this resonance curve at half maximum is a second measure of internal friction. Denoting this width by $\Delta\omega_k$, we have

$$\Delta\omega_k=\eta_k/\omega_k.$$

Since $$\int \mathfrak{U}_k\cdot L\mathfrak{U}_k dv=\omega_k^2,$$

we have, setting $\omega_k=2\pi\nu_k$,

$$\Delta\nu_k/\nu_k=\eta_k/\omega_k^2=Q^{-1}(\nu_k).$$

One favorite method of measuring internal friction is to observe the logarithmic decrement of free oscillation. In order to obtain the relation between this yardstick of internal friction, and the two previous yardsticks, $\Delta\nu_k/\nu_k$ and Q^{-1}, we must solve (1) when $\mathbf{F}$ is zero. Taking the solid to be initially vibrating in the kth normal mode, we set

$$\mathbf{U}(\mathbf{r},\,t)=C_k(t)\mathfrak{U}_k(\mathbf{r})e^{i\omega_k t},$$
$$\mathbf{f}(\mathbf{r},\,t)=C_k(t)\mathfrak{f}_k(\mathbf{r})e^{i\omega_k t}.$$

Neglecting $(d/dt)\log C_k$ in comparison to ω_k, we obtain, after multiplying (1) to the left by $\mathfrak{U}_k$ and integrating over the solid,

$$dC_k/dt=-(\eta_k/2\omega_k)C_k.$$

Comparing the solution of this equation,

$$C_k(t)=C_k(0)\,\exp\,\{-(\eta_k/2\omega_k)t\},$$

with the definition of the logarithmic decrement, Δ_k,

$$C_k(t)=C_k(0)\,\exp\,\{-\Delta_k\nu_k t\},$$

we obtain

$$\Delta_k=\eta_k/(2\nu_k\omega_k)=\pi\Delta\nu_k/\nu_k=\pi Q^{-1}(\nu_k). \qquad (4)$$

Summary

One yardstick of internal friction is defined by (3). For each type of applied force, torsional or tensile, Q is a continuous function of frequency. Two other yardsticks are defined, $\Delta\nu_k/\nu_k$ and Δ_k. These have a meaning only with reference to the natural modes of vibration. They are related to each other, and to the value of Q at the natural frequencies, by Eq. (4).

§3. Calculation of Internal Friction

The equation which governs the transverse vibrations in a reed may be written in a simple form, provided the wave-lengths are long compared to the thickness of the reed. Let the reed lie along the x axis, with its face normal to the y axis. Further let $y(x,\,t)$ denote the displacement function of the reed with respect to its position of rest. Then y is governed by the well-known equation[5]

$$\rho\frac{\partial^2 y}{\delta t^2}+\frac{\delta^2 M}{\delta x^2}=f.$$

Here ρ is the mass per unit length, $M(x,\,t)$ is the bending moment, and $f(x,\,t)$ is the normal component of the applied force per unit length. The bending moment may be calculated directly from the function y, and from a knowledge of the distribution of temperature across the reed. This is done as follows.

We separate M into two parts, M_1 and M_2. The first part is the moment which would arise from the bending of the reed if the temperature gradient across the reed were zero. The second part is the bending moment which arises from the variation of temperature across the reed. The first part is commonly written as

$$M_1=EI\delta^2 y/\partial x^2,$$

where E is Young's modulus, and

$$I=ha^3/12.$$

Here h is the width, and a the thickness, of the reed.

If we let s denote the y coordinate inside the reed, with origin at the center of the reed, and T_0 the average temperature, then

$$M_2=\alpha h\int_{-a/2}^{a/2}(T-T_0)ds. \qquad (5)$$

Here $$\alpha=(\delta X_x/\delta T)_\pi,$$

the suffix π denoting that the normal stresses across the xy and the zx plane are kept constant in the partial differentiation.

In order to calculate M_2 explicitly, we must find the temperature fluctuation $T-T_0$ across

[5] Southwell, *Theory of Elasticity* (Oxford Press, 1936), pp. 172–174.

the reed. Now the time derivative $\delta T/\delta t$ may be written as the sum of two terms. The first is due to the flow of heat by diffusion, and is given by $D\delta^2T/\delta s^2$, where D is the thermal diffusion constant. The second term is due to the generation and absorption of heat by the dilation. Since the stresses Y_y and Z_z are both zero,[6] the strains e_{yy} and e_{zz} are completely determined by e_{xx}. The second part of M_2 may thus be written as $\beta\delta e_{xx}/\delta t$, where

$$\beta=(\delta T/\delta e_{xx})_{Q,\,\pi}.$$

The suffix π has the same significance as before; the suffix Q refers to an adiabatic change of e_{xx}. Now the e_{xx} in a reed is given by $-s\delta^2y/\delta x^2$. Since we need only consider the steady state of forced oscillation, we set

$$y(x,\,t)=Y(x)e^{i\omega t}.$$

We thus obtain the following equation for T:

$$\frac{\delta T}{\delta t}=D\frac{\delta^2T}{\delta s^2}-\left(i\omega\beta\frac{\delta^2Y}{\delta x^2}\right)se^{i\omega t}. \tag{6}$$

The boundary conditions for this equation are obtained by observing that the establishment of temperature equilibrium in the reed is due almost entirely to a flow of heat across the reed, and not to and from the air. Now a zero flow of heat across the surface of the reed implies that the temperature gradient is there zero. Our boundary condition is thus

$$\delta T/\delta s=0 \quad \text{at} \quad s=\pm a/2. \tag{6a}$$

That solution of Eq. (6) which satisfies the condition (6a) may readily be obtained as a trigonometric series. This series will automatically satisfy the boundary condition (6a) if each term satisfies this condition. Such a series is given by

$$T=T_0+\Sigma_k g_k(t)\,\sin\,\{(2k+1)\pi s/a\}. \tag{7}$$

The functions $g_k(t)$ are obtained by substitution of this series into (6), multiplying by $\sin\,\{(2k+1)\pi s/a\}$, and integrating across the reed. We obtain

$$g_k=-\frac{(i\omega\beta\delta^2Y/\delta s^2)s_ke^{i\omega t}}{i\omega+\mu_k}.$$

[6] Reference 5, p. 161.

Here

$$s_k=(2/a)\int_{-a/2}^{a/2}s\,\sin\,\{(2k+1)\pi s/a\}ds. \tag{8}$$

and $$\mu_k=\{(2k+1)\pi/a\}^2D.$$

As our final step in obtaining M_2, we substitute T from (7) into (5). Setting

$$M_2(x,\,t)=\mathfrak{M}_2(x)e^{i\omega t},$$

and separating $\mathfrak{M}_2$ into its real and imaginary parts, we obtain

$$\Re\mathfrak{M}_2=\tfrac{1}{2}\alpha\beta ad\frac{\delta^2Y}{\partial x^2}\Sigma_k\frac{\omega^2}{\omega^2+\mu_k^2}s_k^2,$$

$$\Im\mathfrak{M}_2=\tfrac{1}{2}\alpha\beta ad\frac{\delta^2Y}{\partial x^2}\Sigma_k\frac{\omega\mu_k}{\omega^2+\mu_k^2}s_k^2.$$

The Q of the reed is now, by definition,

$$Q^{-1}=\Im\frac{\partial^2}{\partial x^2}\mathfrak{M}\Big/\Re\frac{\delta^2}{\partial x^2}\mathfrak{M}.$$

This may be written explicitly as

$$Q^{-1}=\frac{(\alpha\beta/E)\Sigma_k\dfrac{\omega\mu_k}{\omega^2+\mu_k^2}(6s_k^2/a)}{1+(\alpha\beta/E)\Sigma_k\dfrac{\omega^2}{\omega^2+\mu_k^2}(6s_k^2/a)}. \tag{9}$$

From (8) we obtain

$$6s_k^2/a=96\pi^{-4}(2k+1)^{-4}.$$

For $k=0$, 1, 2 the right member has the values 0.986, 0.012, 0.0016. Hence very little error is made in replacing $6s_k^2/a$ by unity when $k=0$, by zero for all other values of k.

In the appendix we find that

$$\alpha\beta/E=\tfrac{1}{3}(1-2\sigma)(C_p-C_v)/C_v, \tag{10}$$

where σ is Poisson's ratio. Since the right member of (10) is always less than 0.01, we make only a slight error by neglecting the second term in the denominator of (9). We finally obtain as a very good approximation

$$Q_{\text{th}}^{-1}=\frac{1-2\sigma}{3}\cdot\frac{C_p-C_v}{C_v}\cdot\frac{\omega\mu}{\omega^2+\mu^2}, \tag{11}$$

where $$\mu=(\pi/a)^2D, \tag{12}$$

and where the suffix th is a reminder that the right side of (11) gives only that part of Q^{-1} which is due to the thermodynamical effects above discussed.

§4. Quantitative Discussion

That part of the internal friction which is due to the fluctuation of temperature across the reed is given by Eq. (11). Considered as a function of the angular frequency, ω, Q_{th}^{-1} has a maximum at $\omega=\mu$, where μ^{-1} is the relaxation time for the establishment of temperature equilibrium across the reed, and is given by (12). When $\omega \ll \mu$, the vibration proceeds isothermally, when $\omega \gg \omega$, it proceeds adiabatically. In the transitional region between isothermal and adiabatic vibration, $\omega \sim \mu$, stress and strain are no longer in phase. The hysteresis loop caused by this phase shift gives rise to a so-called "internal friction."

The maximum value of Q_{th}^{-1} is given in Table I for several metals. The last column gives the thickness of the reed, a_{1000}, for which the maximum of Q_{th}^{-1} occurs at a frequency of 1000 cycles sec.$^{-1}$ (expressed in cm).

The *actual* Q^{-1} of a reed will be equal to the sum of the Q_{th}^{-1} given by Eq. (11), and the Q^{-1} which arises from all other causes of internal friction. The *measured* Q^{-1} of a reed will also contain a contribution resulting from energy losses at the supports of the reed, and from acoustical losses. The very careful experiments of Wegel and Walther[1] indicate that the Q^{-1} of metals measured by nearly all previous investigators consisted principally of terms arising from apparatus losses. They found that the Q^{-1} of a metal was a structure sensitive property. Thus (see Table II of reference 1) they obtained values of Q, at 10,000 cycles per second, ranging from 9000 to 52,000 for aluminum, from 500 to 10,000 for copper, from 3800 to 40,000 for steel. These values were obtained for the longitudinal vibrations in cylindrical rods, in which the thermodynamical damping discussed in this paper was not present. If the same materials had been used in the form of a reed, the damping due to causes other than that here discussed would probably have been of the same order of magnitude. Comparing the minimum values of Q^{-1} found by Wegel and Walther for, say, copper and aluminum, 10^{-4} and 0.2×10^{-4}, respectively, with the maximum values of Q_{th}^{-1} in Table I, we find that samples may be found in which Q_{th}^{-1} is at least 10 times as large as the Q^{-1} due to all other causes over a broad frequency band.

Table I. *Values of Q_{th}^{-1} for various metals at 20°C and the thickness (a) of the reed for which the maximum value of Q_{th} occurs at 1000 cycles per sec.*

Metal	$(C_p-C_v)/C_v$*	σ	Max. of $Q_{\text{th}}^{-1}\times10^3$	$a_{1000}\times10^2$
W	0.006	0.17	0.66	3.5
Pl	.020	.387	.76	2.0
Au	.038	.42	.81	4.3
Pb	.067	.45	1.0	2.0
Ni	.021	.31	1.3	1.5
Cu	.028	.34	1.4	4.2
Ag	.040	.38	1.6	5.1
Sn	.040	.33	2.3	2.5
Zn	.052	.33	3.0	2.5
Cd	.060	.30	4.0	2.7

* Eucken, *Handbuch der Experimental Physik* Vol. 8, p. 211.

The formula (11) may be checked both by a variation of frequency, and by a variation of temperature. Temperature has a two-fold effect upon Q_{th}^{-1}. (1) A lowering of the temperature displaces the maximum value of $Q_{\max}^{-1}$ to higher frequencies by increasing the diffusion constant D. (2) From the empirical formula[7]

$$(C_p-C_v)/C_v=\text{const}\times TC_p$$

we see that a lowering of temperature decreases the maximum of Q_{th}^{-1}.

Appendix

The left member of Eq. (10) may be written as

$$\left(\frac{\delta T}{\delta e_{xx}}\right)_{Q,\pi}\left(\frac{\delta X_x}{\delta T}\right)_{e_{xx},\pi}\Big/\left(\frac{\delta X_x}{\delta e_{xx}}\right)_{T,\pi}. \tag{a-1}$$

Here Young's modulus for constant temperature, E, has been written explicitly as $(\delta X_x/\delta e_{xx})_{T,\pi}$. A standard transformation in partial differentiation reduces (a-1) to

$$-\left(\frac{\delta T}{\delta e_{xx}}\right)_{Q,\pi}\left(\frac{\delta e_{xx}}{\delta T}\right)_{X_x,\pi}. \tag{a-2}$$

At this stage it is convenient to replace the strain derivatives by more familiar forms. The temperature

[7] Eucken, *Handbuch der Experimental Physik* Vol. 8, p. 209.

change caused by an adiabatic stretching is equal to $(\delta T/\delta V)_Q \delta V$, where δV is the change in volume accompanying the strain.

Now $e_{yy}=e_{zz}=-e_{xx}$
when $Y_y=Z_z=0.$
Hence $\delta V/V=e_{xx}+e_{yy}+e_{zz}=(1-2\sigma)e_{xx},$
and so $(\delta T/\delta e_{xx})_{Q,\pi}=(1-2\sigma)(\delta T/\delta \log V)_Q.$

The second factor in (a-2) may obviously be written as $\frac{1}{3}(\delta \log V/\partial T)_p$. Hence (a-2) reduces to

$$-\frac{1-2\sigma}{3}\left(\frac{\partial T}{\partial V}\right)_Q\left(\frac{\partial V}{\partial T}\right)_p. \qquad \text{(a-3)}$$

This expression may be written in a more illuminating form. If only two of the four variables x_1, x_2, x_3, x_4 are independent, then the following identity is true:

$$\frac{(\partial x_1/\partial x_2)_{x_4}}{(\partial x_1/\partial x_2)_{x_3}}=1-\frac{(\partial x_2/\partial x_2)_{x_4}}{(\partial x_3/\partial x_2)_{x_1}}.$$

Replacing x_1, x_2, x_3, x_4 by V, T, Q, p, respectively, we find that (a-3) becomes

$$\tfrac{1}{3}(1-2\sigma)(C_p-C_v)/C_p.$$

AUGUST 1, 1937 PHYSICAL REVIEW VOLUME 52

Collector Theory for Ions with Maxwellian and Drift Velocities

A. H. HEATLEY
Electrochemical Laboratory, University of Toronto, Toronto, Canada
(Received February 15, 1937)

For a gaseous discharge in which the ions have a drift velocity superimposed on a Maxwellian distribution, the Mott-Smith and Langmuir collector theory deals only with collectors having a high ratio of sheath to collector diameters. The present paper removes this restriction, but points out that except when there is negligible sheath distortion due to low drift velocity or to low collector potential the theory gives only approximate results. It is also shown that for ratios of drift to random current density as low as seven, the space potential is not indicated by a break on the semi-logarithmic plot of the collector characteristic.

AN electrical discharge in a gas is conveniently studied by the use of a collector or probe connected to a suitable auxiliary circuit. Mott-Smith and Langmuir[1] have shown that if the logarithm of the collector current be plotted against the potential applied to it, the resulting "characteristic" is a line which departs from linearity at the space potential. This holds for a Maxwellian distribution of the ions (or electrons) of the discharge; but when there is a small drift velocity superimposed on the Maxwellian distribution the characteristic is modified as shown in their paper. Their results refer only to the case where there is a high ratio of sheath[2] to collector diameter. The present paper deals with sheaths of all sizes.

The basic assumptions of this collector theory are discussed by Mott-Smith and Langmuir.[1] Special attention is directed to the two assumptions (1) that the gas pressure is so low that collisions between ions or electrons and gas molecules in the sheath have a negligible effect on the collector current, and (2) that there is no reflection of ions or electrons at the collector surface.

Because of the drift velocity, the cross section of the space charge sheath around a cylindrical collector (with its axis at right angles to the direction of the drift) is distorted from the circular shape. But obviously it still must have bilateral symmetry about the direction of the drift, which is therefore taken as the coordinate axis. The length of the collector is assumed to be so great that end effects may be neglected.

Without the restriction imposed by Mott-Smith and Langmuir that the sheath is circular and by methods analogous to theirs, it may be shown that, as a first approximation,

[1] H. M. Mott-Smith and I. Langmuir, Phys. Rev. **28**, 727 (1926).
[2] The sheath boundary may be briefly defined as the surface beyond which the collector potential exerts no force on the ions or electrons of the discharge.

$$\frac{i}{2\pi rlI_r} \approx \frac{2}{\pi^{\frac{1}{2}}r} e^{-\alpha^2} \int_0^{\pi} \int_{0,\ (-\eta\tau/\sigma)^{\frac{1}{2}}}^{\infty} \int_{x\chi/\tau-(\sigma x^2+\tau\eta)^{\frac{1}{2}}/\tau}^{x\chi/\tau+(\sigma x^2+\tau\eta)^{\frac{1}{2}}/\tau} \frac{\partial s}{\partial \psi} x \exp(-x^2-y^2+2\alpha x\cos\psi+2\alpha y\sin\psi)dydxd\psi, \quad (1)$$

where i is the current to the collector, r and l are the radius and exposed length, respectively, of the cylindrical collector, I_r is the random current density corresponding to the Maxwellian distribution, s is the circumference of the sheath around the collector, α is $I_d/2\sqrt{\pi I_r}$ (I_d is the drift current density), η is eV/kT (e is the electronic charge, k is Boltzmann's constant, V is the potential of the collector with reference to the space potential, and T is the temperature corresponding to the Maxwellian distribution), σ is $(b/r)^2-1$, τ is $(b\cos\varphi/r)^2-1$, χ is $(b/r)^2\cos\varphi\sin\varphi$ and φ is $\psi-\theta$. b and θ are the coordinates of points on the circumference of the sheath; and the normal at the point b makes the angle ψ with the coordinate axis. See Fig. 1.

Equation (1) is approximate instead of exact because in its development the angular momentum of an ion at the sheath surface is equated to its angular momentum on arrival at the collector surface; but this equality does not hold for a noncircular sheath since the force on the ion due to the collector potential is noncentral (i.e., it is not directed always to the collector axis). The exact solution of this problem involves the determination of the shape of the sheath and the distribution of the force-field within it, a problem which is not dealt with in this paper. However, the sign of the error can be deduced by introducing the plausible assumption that for retarding potentials on the collector (η negative) the sheath shape is such that ϕ is positive (as in Fig. 1) and that for accelerating potentials (η positive) ϕ is negative. Eq. (1) is equivalent to the assertion that ions will reach the collector if they arrive at the sheath boundary traveling in a certain direction (or within a certain range of directions) and within a certain range of velocities. Owing to the sheath distortion assumed above, the ions which will reach the collector are those having a direction making an angle with the coordinate axis greater than if the sheath was circular; and since the distribution function shows that there are fewer of such ions it follows that the current given by Eq. (1) is too large whether the potential on the collector is retarding or accelerating.

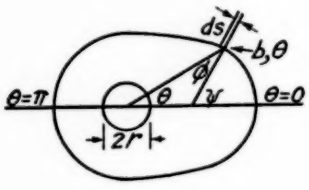


FIG. 1.

For a small sheath, $(b/r=1)$, Eq. (1) becomes exact, since the sheath is now circular. The equation reduces to

$$\mathfrak{I} = \frac{2}{\pi} \int_0^{\pi} \int_{0,\ \sqrt{(-\eta)}}^{\infty} xe^{-(x-\alpha\cos\theta)^2} dxd\theta, \quad (2)$$

where $\mathfrak{I}$ is written for $i/2\pi rlI_r$. This further reduces to

$$\mathfrak{I} = e^{-\alpha^2/2}[(1+\alpha^2)I_0(\alpha^2/2)+\alpha^2 I_1(\alpha^2/2)] \quad \text{for } \eta \geq 0, \quad (2a)$$

$$\mathfrak{I} = e^{\eta-\alpha^2/2}[(1+\alpha^2)I_0(\alpha^2/2)I_0(t)-\alpha^2 I_0(\alpha^2/2)I_2(t) + 2(1+\alpha^2)\sum_1^{\infty}(-1)^p I_p(\alpha^2/2)I_{2p}(t) - \alpha^2\sum_1^{\infty}(-1)^p I_p(\alpha^2/2)\{I_{2p-2}(t)+I_{2p+2}(t)\}] \quad \text{for } \eta \leq 0, \quad (2b)$$

where $t=2\alpha\sqrt{(-\eta)}$ and $I_p(x)$ is the modified Bessel function of the first kind and the pth order.

Without any restriction as to sheath size, but considering the limiting case where the sheath is a circle of radius b, Eq. (1) reduces to

$$\mathfrak{I} = \frac{1}{\pi}\frac{b}{r}\int_0^{\pi}\int_{0,\ \sqrt{(-\eta)}}^{\infty} xe^{-(x-\alpha\cos\theta)^2} \times\left[\frac{2}{\sqrt{\pi}}\,\mathrm{Erf}\,\{[(x^2+\eta)/\sigma]^{\frac{1}{2}}+\alpha\sin\theta\} + \frac{2}{\sqrt{\pi}}\,\mathrm{Erf}\,\{[(x^2+\eta)/\sigma]^{\frac{1}{2}}-\alpha\sin\theta\}\right]dxd\theta, \quad (3)$$

where $\text{Erf}\,(t)=\int_0^t e^{-y^2}dy.$

For a very large sheath $(b/r=\infty)$, Eq. (1) reduces to

$$\mathfrak{I}\approx\frac{4}{\sqrt{\pi}}e^{-\alpha^2}\int_{0,\ \sqrt{(-\eta)}}^{\infty}x(x^2+\eta)^{\frac{1}{2}}e^{-x^2}I_0(2\alpha x)dx, \quad (4)$$

which is equivalent (changing $\approx$ to $=$) to Eq. (46)[3] of Mott-Smith and Langmuir whose basic equation referred to a circular sheath; but it should be observed that Eq. (4) is approximate, not exact, for noncircular sheaths—in disagreement with the statement by Mott-Smith and Langmuir that "the current for very large sheaths must be independent of the actual shape of the sheath."[4] While they reached this conclusion by an alternative method of calculating the collector current, this method also involved the angular momentum of an ion, and is inaccurate for the same reason as Eq. (4).

Since the error in Eq. (4) caused by the noncircularity of the sheath must approach zero as the sheath becomes less distorted, it is justifiable to use it when α is small; but for large values of α when the sheath must certainly be greatly distorted (unless it is small, because of a low value of η) Eq. (4), or any equations such as Eqs. (MS 54) to (MS 57) developed from its equivalent Eq. (MS 46), should not be used until it is shown that their error is negligible.

For negative values of η, Mott-Smith and Langmuir have reduced their Eq. (46) to Eq. (47). It may further be shown that Eq. (2b) is algebraically equivalent to Eq. (MS 47). When $\alpha/\sqrt{(-\eta)}$ exceeds unity, Eq. (2b) is suitable for computation; for low values of the ratio, Eq. (MS 47) is better.[5]

It was found that computations of Eq. (3) for $b/r=\sqrt{2}$ and 10.05 gave the same collector current as obtained from Eq. (2b), $(b/r=1)$ or its equivalent Eq. (MS 47), $(b/r=\infty)$. Thus the collector current is independent of the sheath size for the restricted cases where sheath distortion is negligible, i.e., when α or η are small. Mott-Smith and Langmuir have shown this to be true for three simple types of velocity distributions.[6]

For positive values of η, the solution of Eqs. (3) and (4) depends on a knowledge of the size of the sheath which is not available. Nevertheless, the collector current must lie between the values given by Eqs. (2a) and (4). The latter may be reduced to two equivalent forms,

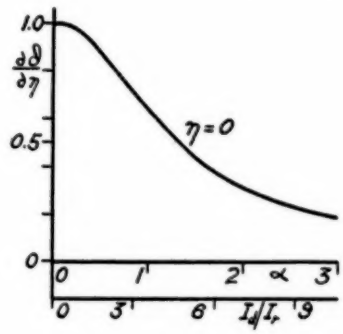


FIG. 2. $\partial\mathfrak{I}/\partial\eta$ as a function of α.

$$\mathfrak{I}\approx e^{\eta-\alpha^2}\sum_0^\infty\frac{(2p+1)!}{(p!)^2 4^p}\left\{(-1)^{p+1}\left(\frac{\sqrt{\eta}}{\alpha}\right)^{p+\frac{1}{2}}J_{p+3/2}(t)+\left(\frac{\alpha}{\sqrt{\eta}}\right)^p J_p(t)\right\}\quad \eta\geq 0, \quad (5a)$$

$$\mathfrak{I}\approx e^{\eta-\alpha^2}\left(\frac{\sqrt{\eta}}{\alpha}\right)^{\frac{1}{2}}\sum_0^\infty(-1)^{p+1}\frac{(2p+1)!}{(p!)^2 4^p}\times\left(\frac{\sqrt{\eta}}{\alpha}\right)^p J_{p+3/2}(t)+e^{\eta-\alpha^2/2}\Big[(1+\alpha^2)I_0(\alpha^2/2)J_0(t)+\alpha^2 I_0(\alpha^2/2)J_2(t)+2(1+\alpha^2)\sum_1^\infty I_p(\alpha^2/2)J_{2p}(t)+\alpha^2\sum_1^\infty I_p(\alpha^2/2)\{J_{2p-2}(t)+J_{2p+2}(t)\}\Big]\quad \eta\geq 0, \quad (5b)$$

where $J_p(x)$ is the Bessel function of the first kind and the pth order, and $t=2\alpha\sqrt{\eta}$. As α and η become large, these series converge very slowly and the computation becomes more laborious than a graphic integration of Eq. (4).

[3] It will be shown below that i_∞ may properly be changed to i in Eq. (MS 46) *et seq.*

[4] Reference 1, p. 752.

[5] Present tables of $I_p(x)$ are very limited. More extensive tables will soon be available. Report B.A.A.S. (1935), p. 304.

[6] Reference 1, p. 751.

When α is zero or very small, the graph of the collector current or its logarithm against collector voltage (i.e., $\eta \times$ a constant) indicates clearly the

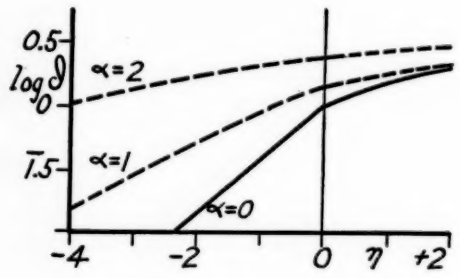


FIG. 3. Log $\mathfrak{Z}$ as a function of η.

space potential. But it should be observed that as α increases, the value of $\partial \mathfrak{Z}/\partial \eta$ (at η equals zero) quickly decreases, (Fig. 2) and therefore the curve of $\mathfrak{Z}$ or its logarithm shows no distinctive change at this point for values of α as low as two. (Fig. 3). Thus this method of determining the space potential cannot be applied in these cases. It should be emphasized that this does not preclude a break at the space potential caused by factors specifically excluded by the basic assumptions of this theory, e.g., reflection of ions when there is a retarding potential on the collector.

The equation for $\partial \mathfrak{Z}/\partial \eta$ can be given in two equivalent forms and the limiting forms at η equals zero are included for completeness.

$$\partial \mathfrak{Z}/\partial \eta = e^{\eta-\alpha^2/2}[I_0(\alpha^2/2)I_0(t) + 2\sum_1^\infty (-1)^p I_p(\alpha^2/2)I_{2p}(t)] \quad \text{for } \eta \leq 0, \quad (6a)$$

$$\partial \mathfrak{Z}/\partial \eta = e^{\eta-\alpha^2}\sum_0^\infty \frac{(2p)!}{(p!)^2 4^p}\left(\frac{\alpha}{\sqrt{(-\eta)}}\right)^p I_p(t) \quad \text{for } \eta \leq 0, \quad (6b)$$

where $t = 2\alpha\sqrt{(-\eta)}$.

$$\partial \mathfrak{Z}/\partial \eta = e^{-\alpha^2/2} I_0(\alpha^2/2) \quad \text{for } \eta = 0, \quad (6c)$$

$$\partial \log \mathfrak{Z}/\partial \eta = 1/[1+\alpha^2\{1+I_1(\alpha^2/2)/I_0(\alpha^2/2)\}] \quad \text{for } \eta = 0. \quad (6d)$$

The limit of Eqs. (2) or (4) for η equals zero, as α increases, is $2\alpha/\sqrt{\pi}$, so that the collector current can be expressed in terms of the drift current instead of the random current.

$$\frac{i}{2rlI_d} = \frac{\pi I_r}{I_d}\cdot \mathfrak{Z} \approx \frac{\sqrt{\pi}}{2\alpha}\cdot\frac{2\alpha}{\sqrt{\pi}} = 1 \text{ (for } \eta = 0). \quad (7)$$

The exact values are 1.282 for $\alpha = 1$, 1.062 for $\alpha = 2$, 1.028 for $\alpha = 3$ (values of $i/2\pi rlI_r$ from Eq. (2a)); so that knowing the drift current density and the collector dimensions we can calculate the minimum value of collector current at the space potential regardless of α.

Attention should be drawn to an improper use of this collector theory by Bramhall,[7] who used a probe in a copper arc. The space potential was determined by a break in the semi-logarithmic plot of the probe characteristic, and the probe current density was there taken to be I_r, the random current density, which is true only when there is no drift current. The drift current density, I_d, was calculated from the arc area and current, giving I_d/I_r roughly 100, or α equals 28. But this high value of α is inconsistent with a break in the characteristic at space potential. Also the probe current calculated from the probe dimensions and the drift current density, Eq. (7), should have a minimum value of 0.4 ampere, whereas the observed value was 0.04 ampere. Further, the part of the theory used by Bramhall in Fig. 9 is based on the Mott-Smith and Langmuir Eq. (54) which is restricted to large values of α, with $\sqrt{-\eta} \approx \alpha$. The value of η in his experiments certainly lay between zero and -25, so that α should not have exceeded 5, a value much too small to warrant the use of Eq. (MS 54)—quite apart from any discussion in this paper as to the validity of that equation.

[7] E. H. Bramhall, Phil. Mag. **13**, 682 (1932).

A Survey of Radioactivity Produced by High Energy Neutron Bombardment

M. L. Pool,* J. M. Cork and R. L. Thornton
Department of Physics, University of Michigan, Ann Arbor, Michigan

(Received July 3, 1937)

A LARGE number of the known stable elements have been bombarded by neutrons of energies ranging up to 20 Mev. The transmutation reaction of the type involving the ejection of two neutrons is apparently evident with varying degrees of intensity in nearly all the elements of the periodic table. In addition, reactions of the type where a proton or an alpha-particle is ejected are observed. For bombarding neutrons of low energy these reactions are observed only for relatively light elements, but for very energetic neutrons such reactions are evidently quite common in elements of high atomic number.

* Elizabeth Clay Howald Scholar.

TABLE I. *Radioactivity produced by very high energy neutron bombardment.*

At. No.	Element bombarded	Period, intensity, sign, assignment	Other work
6	C	20 m, vw, +, C^{11}	
7	N	10.5 m, vw, +, N^{13}	
8	O	2.1 m, vw, +, O^{15}	2.1 m[1]
9	F	108 m, w, +, F^{18} Chem.	
12	Mg	15 h, vs, −, Na^{24}	
13	Al	10 m, vs, −, Mg^{27}	
		15 h, vs, −, Na^{24} Chem.	
14	Si	6 m, vw, +, Si^{27}	
		11 m, vw, +,	
15	P	3 m, vs, +, P^{30}	2.5 m[2]
		2.5 h, s, −, Si^{31}	
16	S	26 m, vw, +, S^{31}	
		14 d, s, −, P^{32}	
17	Cl	33 m, vs, +, Cl^{34} Chem.	
		14 d, s, −, P^{32}	
19	K	7.5 m, s, +, K^{38}	
		4 m, s, −,	
		1.8 h, w, −, A^{41}	
20	Ca	4.5 m, vw, +, Ca^{39}	
21	Sc	4 h, vs, +, Sc^{43}	
		2 d, s, +, Sc^{44}	
22	Ti	1.7 h, vs, −, Sc	
		28 h, s, −, Sc	
23	V	4 m, s, −, V^{52} Chem.	
		1.8 d, vw,	
24	Cr	4 m, vs, −, V^{52}	
		1.7 h, s, −,	
25	Mn	3.6 m, vs, −, V	
		2.5 h, s, −, Mn^{56}	
26	Fe	2.5 h, vs, −, Mn^{56}	
27	Co	2.5 h, s, −, Mn^{56} Chem.	11 m[3]
28	Ni	2 h, w,	2.7 h[3]
		6 d, vw,	11 m[3]
29	Cu	10 m, vs, −, Cu^{62}	10.5 m[2, 4, 3, 1]
		12.5 d, vs, ±, Cu^{64} Chem.	
30	Zn	6 m, vs, −, Cu^{66}	6 m[4]
		40 m, vs, +, Zn Chem.	38 m[2], 60 m[4, 3]
		12.5 h, vs, ±, Cu^{64} Chem.	
31	Ga	20 m, vs, −, Ga^{70}	20 m[2], 5 m[1]
		55 m, vs, +, Ga^{68} Chem.	60 m[1]
		1.7 h, s, +,	
		22 h, w, −, Ga^{72}	
32	Ge	1.3 h, s, −,	
		20 h, w, −, Ge^{72}	
33	As	1.1 d, vs, −, As^{76} Chem.	
		13 d, w,	
34	Se	1 h, vs, −, Se Chem.	56 m[4]
35	Br	7 m, vs, , Br^{78}	3.5 m[2], 5 m[4]
		18 m, vs, , Br^{80} Chem.	18 m[4, 2, 5]
		4 h, s, , Br^{80} Chem.	5 h[1], 24 h[5]
37	Rb	11 m, vs, −,	
		22 h, s, −,	
38	Sr	18 m, s, −,	
		3 h, s, +, Sr Chem.	
39	Y	11 m, s, −,	
		1.2 h, s,	
		6.5 h, w, −,	
40	Zr	2.4 d, vw, −, Y^{90}	
		10 m, s,	
		5 h, w, , Zr	
		44 h, w, , Zr	
41	Cb	7.3 m, w,	
		3.8 d, w,	
42	Mo	17 m, s, −,	21 m[4], 17 m[2]
		5 d, w, −,	
44	Ru	24 m, w,	
		3.6 h, w,	
45	Rh	4 m, s, −,	
		1.1 h, w, −,	
46	Pd	18 m, w, −, Pd	
		12.5 h, w, −, Pd	
47	Ag	25.5 m, vs, +, Ag^{106} Chem.	24 m[4, 2, 1]
		13 d, w, −, Ag Chem.	
48	Cd	33 m, s, +,	
		3 h, w, −, Cd^{115}	3.3 h[4]
		53 h, vw, −, Cd^{117}	
49	In	1.1 m, s, −,	1 m[2, 1]
		54 m, s, −, In Chem.	
		4 h, vs, −, In Chem.	
		2 mo, s, −, In Chem.	
50	Sn	47 m, s, −,	
51	Sb	15.4 m, s, +, Sb^{120}	13 m[2, 1] 17 m[4]
		2.3 d, w, −, Sb^{122}	
52	Te	1.1 h, vs, −,	60 m[4, 2]
		30 d, w,	
53	I	26 m, s, , I^{128} Chem.	
56	Ba	2.5 m, s,	
		85 m, w, , Ba^{139}	
57	La	2.2 h, w,	
58	Ce	40 m, w, −,	
59	Pr	3 m, w, +, Pr^{140}	
		20 h, w, , Pr^{142}	
60	Nd	2.2 h, w, −,	
		19 h, w, −, Pr^{142}	
64	Gd	19 h, w, −,	
66	Dy	2.5 h, w, −,	
73	Ta	9.1 h, w,	14 m[2]
75	Re	18 h, s, −, Re^{186}	
77	Ir	15 h, s, −, Ir^{192}	
78	Pt	1.8 h, s, −,	
		3 d, vw,	
79	Au	17 m, w,	
		2.5 d, w, , Au^{198}	
80	Hg	45 m, s, −,	43 m[4]
81	Tl	5 m, s, , Tl^{200}	4.1 m[4]
		50 m, vw, , Tl^{204}	
82	Pb	5 m, w, −,	
		1.5 h, w, −,	
90	Th	5 m, vs, −,	
		1.4 h, s,	
92	U	26 m, vs, −,	
		4 h, s, −,	
		13 h, s, −, Eka Os	

[1] Chang, Goldhaber and Sagane, Nature **138**, 962 (1937).
[2] Bothe and Gentner, Naturwiss. **25**, 90, 126, 191 (1937).
[3] Heyn, Physica **4**, 160 (1937).
[4] Heyn, Nature **138**, 723, 842 (1937).
[5] Johnson and Hamblin, Nature **138**, 504 (1936).

These neutrons of high energy are obtained by the previously reported method of bombarding lithium in the cyclotron with 6.3 Mev deuterons.[1]

The results are summarized in Table I. The assignment of the periods is tentative and is based upon evidence from the sign of the emitted beta-particle, the chemical separations and known periods from other sources. Detailed work is necessary in many cases to make identification positive. Of the 113 periods listed in the table only about a fourth can also be obtained by slow neutron bombardment. Bismuth was the only element tried that gave no measurable radioactivity.

The authors are very grateful for help from various sources. Professor Quill of the Ohio State University supplied very pure samples of several of the more rare elements. Professor Hopkins of the University of Illinois loaned the dysprosium, previously used for atomic weight measurements. Mr. D. W. Stewart, Mr. E. Rosenbaum and Miss G. Mueller made the chemical separations. The aid of Mr. B. R. Curtis in connection with the cloud chamber observations is also gratefully acknowledged. This work was made possible by a grant from the Horace H. Rackham fund.

[1] Pool, Cork and Thornton, Phys. Rev. **51**, 890 (1937).

Important Notice

THE American Institute of Physics will pay $2.00 per copy for each of the following issues of the *Physical Review:*

Volume 1, No. 6—June, 1913
Volume 2, No. 2—August, 1913
Volume 2, No. 4—October, 1913
Volume 2, No. 5—November, 1913
Volume 7, No. 1—January, 1916
Volume 7, No. 2—February, 1916
Volume 7, No. 3—March, 1916
Volume 7, No. 4—April, 1916
Volume 7, No. 5—May, 1916
Volume 9, No. 2—February, 1917

Copies should be in good condition. They should be addressed to the Publications Manager, American Institute of Physics, 175 Fifth Avenue, New York, New York, with the name and address of sender and with a covering letter.

LETTERS TO THE EDITOR

Prompt publication of brief reports of important discoveries in physics may be secured by addressing them to this department. Closing dates for this department are, for the first issue of the month, the eighteenth of the preceding month, for the second issue, the third of the month. Because of the late closing dates for the section no proof can be shown to authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents.

Communications should not in general exceed 600 words in length.

On the Values of Fundamental Atomic Constants

A recent article by von Friesen,[1] with the above title, seems to call for certain comments. In the first place von Friesen writes his adopted values in the form $x\pm t$, where $2t$ indicates "the narrowest region within which one is sure to find the true value." He also argues that the concept of probable error can be applied only to accidental errors.

In my article on the general physical constants,[2] I gave reasons why it is preferable to publish the value of a physical constant with its estimated probable error, rather than with an estimated "limit of error." The situation is essentially as follows. When a person states that he is "sure" that the value of a physical constant lies within a certain region, he does not really mean this. One can never be absolutely sure. What one has in mind is that it seems reasonable to bet some very high odds, such as 100 to 1, or 1000 to 1, that the value in question lies within the stated limits. Now the concept of "probable error" can equally well be applied in any such situation, since it merely defines a region for which the odds are *even* that the true value lies within it. It is entirely unnecessary to base the estimated probable error merely on accidental errors. In fact one of my main efforts during the past ten years has been to emphasize the necessity of considering also systematic errors in any published estimate of probable error. Hence when I write $x\pm r$ for a physical quantity, I mean merely that, in the light of *all* the facts available to me, it appears an even bet that the true value lies within an interval $2r$, centered on x.

In considering experimental values of physical constants, it is not unreasonable to assume a normal distribution of errors, since in general actual distributions found in practice seem to approximate the normal error curve more closely than any other convenient mathematical distribution. Now for a normal distribution there is one chance in 100 of getting an error greater than $3.82r$ (where r is probable error), and 1 in 1000 of exceeding $4.90r$. Hence if a writer chooses to give what he terms a "limit of error," it is necessary merely to divide this result by, let us say, 4 or 5, to get a corresponding estimate of probable error. It is obvious that the reverse process is equally possible. It would, however, be very advantageous to have a uniform procedure, and mainly for historical reasons I have advocated, and still advocate, the publication of values of constants with an estimate of the region of the "even bet" (i.e., probable error), rather than of the region corresponding to other arbitrarily adopted odds.

The experimental data given by von Friesen for each of the constants e, h and e/m appear very consistent, but a fundamental discrepancy, which he fails to mention, exists between his finally adopted values of these three constants. These values are $e=(4.800\pm0.005)\times10^{-10}$ e.s.u., $h=(6.610\pm0.015)\times10^{-27}$ erg·sec., $e/m=(1.7585\pm0.002)\times10^{7}$ e.m.u., where his uncertainties, as noted, are intended to represent "limits of error." Now these three adopted values do *not* satisfy the Bohr formula for the Rydberg constant, a formula which von Friesen accepts as correct. In fact, with his adopted values of e and h, one gets, from this formula, $e/m=1.766$, a value far outside his stated limit of error.

The very disconcerting discrepancy that exists between the directly measured values of e, h/e, and e/m has already been emphasized.[3] The experimental evidence that has since appeared has served only to sharpen the inconsistency in question. I have discussed this situation with many persons and it is the present consensus of opinion that the "grating value" of e is substantially correct (I gave 4.8029 ± 0.0005 for this, in 1936), and that the Bohr formula for the Rydberg constant should be at least very closely correct. When one substitutes in this formula the value of e just quoted, and von Friesen's adopted value of e/m (1.7585), which is certainly close to the present best observed value, the resulting value of h/e is 1.3796×10^{-17} e.s.u., and this is *greater* than *any* directly observed value. The recent beautiful work on h/e by DuMond and Bollman[4] brings into even greater prominence the serious discrepancy between the precise directly observed values of h/e, and the value calculated indirectly from e and e/m.

In spite of the excellence of the recent experimental work on h/e, of which no serious criticism has yet been voiced, I have, in common with many others, reluctantly been forced to the conclusion that *all* of the present experimental values of h/e are low, and that the true value will eventually be found to be in the neighborhood of that just calculated. With $e=4.8029$ this value requires that $h=6.626\times10^{-27}$ erg·sec., and that $1/\alpha=137.044$, in close agreement with Eddington's predicted value of 137. It is, however, only fair to admit that, until a reasonable source of error has been suggested for the present experimental values of h/e, the true values of h and of α remain very uncertain.

RAYMOND T. BIRGE

University of California,
Berkeley, California,
July 13, 1937

[1] Sten von Friesen, Proc. Roy. Soc. **A160**, 424 (1937).
[2] R. T. Birge, Rev. Mod. Phys. **1**, 1 (1929).
[3] R. T. Birge, Nature **137**, 187 (1936).
[4] J. DuMond and V. Bollman, Phys. Rev. **51**, 400 (1937).

On the Capture of Thermal Neutrons by Deuterons

Recent experiments[1] on the radiative capture of thermal neutrons by deuterons indicate as an upper limit for the cross section 3×10^{-26} cm^2. The calculations for this process are readily carried through with the results of a recent paper[2] on the elastic scattering of neutrons by deuterons. The notation, coordinate system, and numerical values used here are the same as those used in reference 2.

The initial state (including the exchange wave) antisymmetric in the neutrons (particles 1 and 2) is

$$\psi^i=(1/\sqrt{2})[\chi(r_{13})\phi_0(r_{23})S(123)-\chi(r_{23})\phi_0(r_{13})S(213)];$$
$$\chi(r)=1+(\Delta/r)(1-e^{-\epsilon_0 r}); \quad \phi_0(r)=(a^3/\pi)^{\frac{1}{2}}e^{-ar};$$

where the S's are the quartet and doublet spin functions given in Eq. (19) of reference 2. We have (in nuclear units) $a=3.236$, $\epsilon_0=5.701$, $\Delta=-0.916$ for the quartet state, and $\Delta=-0.381$ for the doublet state. The final H^3 bound states are

$$\psi_{\pm\frac{1}{2}}{}^b=uR_{\pm\frac{1}{2}}+vR'_{\pm\frac{1}{2}}; \quad R_{\frac{1}{2}}=(1/\sqrt{2})[(+-+)-(-++)];$$
$$R_{-\frac{1}{2}}=(1/\sqrt{2})[(+--)-(-+-)];$$
$$R_{\frac{1}{2}}'=(1/\sqrt{6})[(+-+)+(-++)-2(++-)];$$
$$R'_{-\frac{1}{2}}=(1/\sqrt{6})[2(--+)-(-+-)-(+--)];$$

the space functions u and v are, respectively, symmetric and antisymmetric in the neutrons. The initial quartet and doublet states can be written

$$\psi_q{}^i=(1/\sqrt{2})I^-S_q(123); \quad \psi_d{}^i=(1/2\sqrt{2})[\sqrt{3}I^+R_{\frac{1}{2}}-I^-R_{\frac{1}{2}}'];$$
$$I^\pm=\chi(r_{13})\phi_0(r_{23})\pm\chi(r_{23})\phi_0(r_{13}).$$

The orthogonality condition requires that $\sqrt{3}\ (u,I^+)-(v,I^-)=0$; also $(u,I^-)=(v,I^+)=0$.

At thermal energies, the transition occurs under the influence of the magnetic dipole $\mathbf{G}=g_n(\boldsymbol{\sigma}_1+\boldsymbol{\sigma}_2)+g_p\boldsymbol{\sigma}_3$. The mean transition probability is given in terms of

$$\bar{G}^2=\tfrac{1}{3}\sum_m\sum_{xyz}\{2|(\psi_m{}^b, G_x\psi_q{}^i)|^2+|(\psi_m{}^b, G_x\psi_d{}^i)|^2\},$$

and the inverse mean life of a thermal neutron before capture is[3]

$$(1/\tau)=n(m_e/m_p)^{3/2}(W_\tau{}^3e^2h/6\pi m_p{}^2c^2)\bar{G}^2;$$

here, n is the number of deuterons per cc of absorber, W_τ is the energy in nuclear units (12.2 m_ec^2) given up to the photon, and $\bar{G}^2$ is in nuclear units. With the help of the orthogonality condition, it can be shown that

$$\bar{G}^2=(1/18)(g_n-g_p)^2\{2|(v,I_q{}^-)|^2+|(v,I_d{}^-)|^2\}.$$

We take for u the normalized function of the form[4] $e^{-\mu(r_{13}+r_{23}+\sigma r_{12})}$ with $\mu=2.604$ and $\sigma=0.8$. To estimate v, which is probably small, we write it as $v=(K\mu/2)(r_{13}-r_{23})u$, where K is a numerical factor. The calculations of Rarita and Present[5] indicate that $K=0.013$; although this value certainly will not be exact for our case, it serves to give an order of magnitude.

The integrals (v,I^-) can be estimated; with $g_n=-4.0$ and $g_p=5.7$ we obtain $(1/\tau)\leq50$ sec.$^{-1}$ for D_2O ($n=6.7\times10^{22}$ deuterons per cc). This estimate is probably an upper limit to the inverse mean life. For thermal neutrons (mean velocity 250,000 cm/sec.), this corresponds to a cross section for capture of 0.3×10^{-26} cm^2. This is well within the observed[1] upper limit of 3×10^{-26} cm^2, and indicates that the capture process is likely to be too improbable to be observed experimentally.

I am indebted to Professor G. Breit for valuable discussion of the properties of the spin transitions. These calculations arose through discussion with Professor Philip M. Morse, whom I wish to thank.

L. I. Schiff

Massachusetts Institute of Technology,
Cambridge, Massachusetts,
June 7, 1937.

[1] Kikuchi, Aoki and Takeda, Tokyo Institute of Physical and Chemical Research **31**, 195 (1937).
[2] Schiff, Phys. Rev. **52**, 149 (1937).
[3] Morse, Fisk and Schiff, Phys. Rev. **50**, 748 (1936).
[4] Fisk, Schiff and Shockley, Phys. Rev. **50**, 1090 and 1191 (1936).
[5] Rarita and Present, Phys. Rev. **51**, 788 (1937).

Erratum: Some Lattice Sums Involved in the Calculation of Elastic Constants

(Phys. Rev. **50**, 99 (1936))

In our letter of May 26, 1936 entitled "Some Lattice Sums Involved in the Calculation of Elastic Constants" an error occurred in the calculation of the electrostatic part of C_{12} for the CsCl type lattice from C_{11} and Madelung's constant. The contribution to C_{11} remains unchanged

$$A^{(e)}=(e^2/\delta^4)\times2.1253$$

whereas the revised value of C_{12} is

$$B^{(e)}=-(e^2/\delta^4)\times2.0803.$$

M. Goeppert-Mayer

The Johns Hopkins University,
Baltimore, Maryland.

Albert May

The Catholic University,
Washington, D. C.,
June 3, 1937.

Proceedings of the American Physical Society

Minutes of the Madison, Wisconsin Meeting, June 22–23, 1937

THE 214th regular meeting of the American Physical Society was held at Madison, Wisconsin on Tuesday and Wednesday, June 22 and 23, 1937 at the University of Wisconsin. All sessions were held in the physical laboratory building, Sterling Hall. There was one session for contributed papers on Tuesday morning and a symposium on Tuesday afternoon. On Wednesday morning there were two parallel sessions for contributed papers. The presiding officers were H. M. Randall, President of the Society, and John R. Roebuck. The visiting members of the Society were housed in Chadbourne Hall. The attendance at the meeting was about one hundred and fifty.

A symposium on "Generation and Studies of High Frequency Currents of Centimeter Range" was held on Tuesday afternoon at two o'clock in Sterling Hall with H. M. Randall presiding. The invited papers were as follows: (1) *The Production and Absorption of Electromagnetic Waves Having Length from 2 cm to 6 mm* by Neil H. Williams, University of Michigan; (2) *Some New Experimental Methods Applicable to Extremely High Frequencies* by G. C. Southworth, Bell Telephone Laboratories Inc.; (3) *The Magnetron as a High Frequency Generator* by G. R. Kilgore, RCA Manufacturing Company; and (4) *The Negative Grid Tube as a High Frequency Generator* by A. L. Samuel, Bell Telephone Laboratories Inc. The attendance at this symposium was about one hundred and thirty-five.

The local committee had arranged for a boat ride on Lake Mendota, a visit to the United States Forest Products Laboratory and a picnic on Observatory Hill. On Tuesday afternoon Mrs. C. E. Mendenhall served tea to those in attendance at the meeting.

The dinner was held in Tripp Commons, Memorial Union at seven o'clock on Tuesday evening. President Randall presided and called upon President Dykstra of the University of Wisconsin, John T. Tate and J. H. Van Vleck. There were one hundred and sixty-seven guests at the dinner.

Meeting of the Council. At its meeting held on Tuesday, June 22, 1937 at twelve-thirty o'clock the deaths of one fellow (Frederic E. Ives) and two members (Kenneth Hartley and Archer Hoyt) were reported. Three candidates were elected to fellowship, nine candidates were transferred from membership to fellowship and fifty-three candidates were elected to membership. *Elected to fellowship:* Felix Bloch, G. Gamow and Harry J. White. *Transferred from membership to fellowship:* Sidney W. Barnes, Preston R. Bassett, Richard A. Beth, Emma P. Carr, Carl T. Chase, Arthur E. Haas, Milton S. Plesset, Frederick Seitz and Alfred Wolf. *Elected to membership:* Philip Abelson, Reina Albagli, Kyozo Ariyama, Yardley Beers, Alva H. Bennett, Walter Burg, Edward C. Campbell, John S. Campbell, R. H. Cole, Richard C. Davis, Marc de Hemptinne, Gerhard Dessauer, J. F. Eichelberger, Milton E. Gardner, Malcolm W. Gay, Bernard Goldberg, M. Goldhaber, Sterling Gorrill, Archie C. Grant, D. V. Guthrie, Allen V. Hershey, Theodore Holstein, Tadao Horie, Juro Horiuti, William Huppert, David C. Kalbfell, Irving Katz, T. J. Keary, Joseph M. Keller, Arthur F. Kip, Gertrud Kornfeld, Hoff Lu, Richard C. McCurdy, Austin C. McTigue, Robert Marshak, James F. Marvin, Ralph D. Myers, Atherton Noyes, Jr., E. H. Plesset, Alexander Rusterholz, Robert K. Saxe, W. E. Shoupp, Paul L. Smith, H. S. Sommers, Jr., F. A. Steele, John J. Turin, George M. Volkoff, Vera A. Walchuk, Katharine Way, John U. White, James R. Wilson, Robert W. Wilson and Stanley W. Woodard.

The regular scientific program of the Society consisted of thirty-six contributed papers of which numbers 1, 2, 5, 16, 22, 24, 25 and 27 were read by title. The abstracts of these papers are given in the following pages. An Author Index will be found at the end.

W. L. Severinghaus, *Secretary*

Abstracts

1. Acoustical Detection of Purely Mechanical Vibrations in Quartz Plates. W. H. Pielemeier, *Pennsylvania State College.*—Correlation of extensive results obtained from Y-cut quartz plates with those obtained from X-cut and other plates leads to the conclusion that certain thickness vibrations may produce a considerable acoustical effect while their electrical effect may be absent or difficult to detect by ordinary methods. In order to be maintained such vibrations must be mechanically coupled with one of the piezoelectric modes of the plate having nearly the same frequency. The latter may be a shear or a longitudinal vibration. Obviously thickness vibrations are most effective in producing sound waves emanating from the largest face of the plate. If these are purely mechanical an electrical engineer would probably not suspect their existence. A simple method to obtain their frequency is to measure the wave-length that they produce in a gas in which the supersonic velocity is known. As the most convenient gas is air very careful velocity measurements were made with *known* frequencies in air at barometric pressure, at constant temperature and over a range of humidities. Two anomalous humidity regions exhibiting dispersion were discovered. These partly account for the rather *large* disagreements amongst the velocity measurements made by many reliable investigators.

2. Supersonic Dispersion in Air. W. H. Pielemeier, *Pennsylvania State College.*—The above cited anomalous humidity regions were found at 30°C and a frequency of 528.4 kc/sec. Scattered measurements were also made at other temperatures and frequencies. Knudsen's[1] curve relating critical absorption frequency and H_2O concentration in oxygen was extrapolated to this frequency. One of the above humidity regions was found at 45 percent of the extrapolated critical H_2O concentration for oxygen at 528.4 kc/sec. For lower frequencies Knudsen[1] reported approximately half as much H_2O concentration for air as for O_2. Sinnes and Roseveare[2] show experimentally that in O_2 a velocity dispersion accompanies an absorption peak as expected from theory. The observed velocity shift in this region fits that calculated for the O_2 concentration in air. The shift in the other observed anomalous region fits that for approximately the CO_2 concentration in air reported by Reid.[3] The location of this region is in harmony with the measurements of other investigators.

[1] V. O. Knudsen and L. Obert, J. Acous. Soc. Am. **7**, 249 (1936).
[2] L. S. Sinnes and W. E. Roseveare, J. Chem. Phys. **4**, 427 (1936).
[3] C. D. Reid, Phys. Rev. **35**, 814 (1930).

3. Alternating Current Impedance of *Nitella*. H. J. Curtis and K. S. Cole, *Columbia University.*—The fresh water plant *Nitella* has single cells 10 to 20 cm long and about 0.5 mm in diameter. A single cell may be placed in various electrolytes between two electrodes parallel to the cell axis and the alternating current impedance measured at frequencies from 30 to 2,500,000 cycles per second. The data indicate that the thin cellulose envelope has a specific resistance of about 1000 ohm cm which is nearly independent of the conductivity of the external electrolyte. Inside of this cellulose is the physiological membrane which is found to be virtually nonconducting and to have a capacity of 0.9 to 1.0μf/cm^2 at 5000 cycles with a phase angle of 80°. The specific resistance of the sap, which nearly fills the interior, may be calculated from the measurements on the living cells, but such values disagree sharply with the 60 ohm cm obtained on extracted sap. The difference may be due to the presence, in the living cell, of the chloroplasts which seem to have a complicated electrical structure.

4. A Thermo-Stromuhr for Measuring Blood Flow. J. F. Herrick and E. J. Baldes, *Divisions of Experimental Medicine and Biophysics, The Mayo Foundation, Rochester, Minnesota.*—The physiological processes of bodily tissues depend upon an adequate blood supply. An accurate method of measuring blood flow to a given region or organ offers a means of studying some of these physiologic processes. Also, the heat-regulatory mechanism and the blood-pressure regulatory mechanism depend, to a large extent, on blood flow. To obtain a method which measures the flow of blood without the disturbing factor of anesthesia and without interfering with the normal relationships in the intact dog is the ultimate goal of any investigator in this field. In 1928, Rein[1, 2] described an ingenious indirect method for measuring blood flow in the anesthetized dog. We have been using this method for several years and have applied it to the intact animal. The purpose of the present paper is to report an analysis of physical principles involved in this method. We consider the following: (1) Certain criticisms of Rein's description of the nature of heat transfer; (2) discrepancies of the method when measuring small flows; (3) limitations of the method; (4) a modified thermostromuhr; (5) dependence of the method on the character of the flow.

[1] H. Rein, Zeits. f. Biologie **87**, 394–418 (1928).
[2] H. Rein, Abderhalden Handbuch der biol. arbeit. **8**, 693–716 (1928–35).

5. Thin-Film Lubrication of Non-planar Surfaces of Finite Width. Selby M. Skinner, *Columbia University.*—The classical treatment of thin film lubrication by Sommerfeld is restricted to surfaces of infinite width, thus neglecting escape of the lubricant along the lateral edges of the bearing. This leakage reduces materially the total load which can be supported by the bearing. The plane bearing of finite width has been treated by Michell and others, leading to the development of the Michell thrust bearing. In practice, however, the bearing surface may be curved. The theory is here extended to two general classes of curved bearing surfaces of finite width. A general method of solution is given, and the exact solutions for several cases are developed. The effect of side leakage,

the total load which can be supported, the frictional resistance, and the total quantity of lubricant required are determined. The theory for sector-shaped surfaces is also extended to a case in which the dependence of film thickness on distance from the center of rotation may be treated.

6. Determination of Fiber to Void Ratio of Wood Pulp Paper by Electrical Methods. E. W. GREENFIELD, *Anaconda Wire & Cable Co., Hastings-on-Hudson, New York.*—Application of the method of mixtures has been attempted for determining, by electrical means, the ratio of fiber to voids in a number of commercial wood pulp cable papers. This problem has its important bearing on the paper industries' partial fraction measurements as well as on the oil to paper ratios of impregnated paper high-voltage cables. Separate measurements, under uniform conditions, of the dielectric constants of dry paper, oil and impregnated paper yield sufficient experimental data to enable the calculation of fiber to oil ratio as well as the dielectric constant of the fiber itself. The principal difficulty in using this method is the necessity of an assumed geometric distribution of the cellulose fibers and the voids, and this is of course unknown. Previous attempts* to use this method assumed, for simplicity of calculation, a two-layer series dielectric. The present paper gives results using four different distributions of fiber and void: series layers, parallel layers, series-parallel layers, and cylinders of voids in a medium of solid cellulose perpendicular to the electric field. Correlation of the values obtained for each dielectric distribution with experimental extraction and theoretical density determinations yields interesting data on each paper studied.

* J. B. Whitehead, Trans. A. I. E. E. **2**, 667 (1933).

7. Dynamic Characteristic Curves of Glow-discharge Tubes. HERBERT J. REICH AND WALLACE A. DEPP, *University of Illinois.*—The dynamic current-voltage curves of glow-discharge tubes are observed by means of a cathode-ray oscillograph. The technique differs from that ordinarily used in that the manner in which the current varies with time is under control. This is accomplished by the use of a vacuum tube to limit the current passed by the discharge tube. The rate at which the current rises and falls, the frequency of repetition of the current cycle, and the interval between cycles can be varied at will by changing the waveform of the voltage applied to the grid of the control tube. The marked departure of many of the characteristic curves from the form commonly assumed is explainable on the basis of the time taken for ionization and deionization of the gas.

8. High Pressure Arc Phenomena. C. G. SUITS, *General Electric Company.*—A comprehensive investigation of the properties of the electric arc in the high pressure range has been partially completed. Several pressure chambers and some new measurement methods allow observations of total arc voltage e as a function of current i up to 3600 atmospheres; e, electric gradient E (volts cm^{-1}), current density I (amp. cm^{-2}) assumed constant over the arc cross-section, and light output up to 220 atm. In gases of 99 percent purity the measurements extend, in N_2 from 1–1200 atm., H_2 from .01–200 atm., A from 1–100 atm., and He from 1–200 atm. for Cu, W, and C electrodes. In all gases e, E, and I increase with pressure; at constant pressure e, E, and I increase with gas in the order Hg, A, N_2, He, H_2. The pressure effect is such that e increases approximately 2 times, E 3 times, and I 15 times when p increases 100 times. Since $(n^2/N)\sim(I/E)^2$ for each gas, the gas temperature can be calculated from the Saha equation for an assumed ionization potential. At 1 atm. and 10 amp., the following temperatures, °K, are obtained: A, 6500; He, 5900; Hg, 5400; N_2, 6100 (measured by sound velocity); and N_2 at 100 atm., 7100. The dependence of arc temperature on current and arc length is known to be small from sound velocity measurements. The arc temperature is thus relatively constant over a large range of experimental conditions. The interpretation of the measurements in terms of conduction and convection heat loss from cylinders will be treated in a paper now in preparation. Because the arc temperature is relatively constant, a generalization which predicts the trends with pressure and gas is: *The arc properties E and I are the same in various gases at the pressure at which the conduction and convection heat loss is the same.* Because the temperature is not perfectly constant, some anomalies occur. The measurements of I, and especially E, are subject to considerable uncertainty due to experimental error and lack of reproducibility of experimental conditions, and are to be regarded as tentative in absolute magnitude but correct in trend. Initial data have been obtained with a new method of simultaneously measuring e, I, and E as $f(i)$ with a relatively high accuracy.

9. Photoelectric Sensitization of Aluminum. LOUIS P. THEIN, S.M., *St. Louis University. (Introduced by J. J. Brady.)*—It is well known that many metal surfaces become more sensitive photoelectrically after exposure to a hydrogen glow discharge. In the present work, it is shown that no sensitization occurs if the aluminum target is made the *anode* of the discharge. The increase in photo-current, with the target as cathode, is found to be dependent on the time of the glow discharge and the pressure of the hydrogen. If the target is not outgassed, a rapid decay in photo-current is noted after sensitization, because of the diffusion to the surface of gas contained in the body of the metal. Results indicate that this sensitization is due mainly to a "clean up" of the surface of the oxide-coated aluminum, and the upper limit of this type of sensitization is fixed by the yield of pure aluminum. Another type of sensitization occurs when small amounts of dry air are admitted in contact with freshly distilled aluminum. The sensitivity here is found to be as much as 6 or 7 times as high as that of pure aluminum or as much as 800 times that of an untreated piece of aluminum. The current-voltage data from freshly distilled aluminum yield a curve which fits well upon the theoretical curve of the DuBridge theory. The curves from the surfaces sensitized with air show no agreement with the theory. The threshold for pure aluminum is at 2900A.

10. Temperature Effects in Photoelectric Emission. R. J. CASHMAN, *Northwestern University*.—Fowler's simple photoelectric theory gives for the spectral distribution function $I=CT^2\varphi[(h\nu-h\nu_0)/(kT)]$. For certain values of ν the effect of the T term is much more pronounced for emission at different temperatures than the φ term, a situation unlike that in Richardson's equation for thermionic emission. At $\nu=\nu_0$, $I=C\pi^2/12T^2$. All published data on clean metals show that the variation of I with T is smaller than that predicted by a T^2 law, the value of the exponent changing from about 1 to 1.6. Although one must not overlook the possibilities of a transmission probability which would modify the theory it seems at present that the apparent departure from a T^2 law can be adequately accounted for by introducing a small positive temperature coefficient of the work function. Let $h\nu_0$ and $h\nu_0+\Delta w$ be the work functions at T_0 and T, respectively. The currents at $h\nu=h\nu_0$ will be, according to Fowler's theory,

$$I_0=C\pi^2/12T_0^2, \qquad I=CT^2\varphi(\Delta w/kT),$$

where φ is a universal function of the argument. Values for the coefficient $\Delta w/(T-T_0)$ from 3 to 10×10^{-5} ev/deg. will account for the apparent departure of existing data. A positive temperature coefficient also predicts that the current-frequency curves for higher temperatures will cross those for lower temperatures for $\nu\gg\nu_0$, an experimental fact which caused considerable controversy when Fowler's theory was proposed.

11. Quantum Mechanical Calculation of the Lattice Energy of NaCl. ROLF LANDSHOFF, *University of Minnesota*.—Starting from a wave function in the form of a determinant of single-electron wave functions, the lattice energy of the NaCl crystal, consisting mainly of the electrostatic and the exchange energy, is calculated numerically. For single-electron wave functions we take the orbitals at the respective ions, which have been calculated by Fock and Hartree.[1, 2] These orbitals take into account the exchange interaction between the electrons in a single ion. This gives a considerable improvement over the calculation[3] based on orbitals of the original Hartree type in which this exchange is neglected.[4] An attempt to calculate the polarization energy along the lines of Schroedinger's second order perturbation theory yielded values much too large. Therefore only the van der Waals part of the second order energy was included in the calculation. The results are: lattice energy 182 kcal./mole, lattice distance 2.78A and compressibility $4.35\cdot10^{-12}$ bar^{-1}.

[1] V. Fock and Mary J. Petrashen, Zeits. d. Sowjetunion **6**, 368 (1934).
[2] D. R. Hartree and W. Hartree, Proc. Roy. Soc. **A156**, 45 (1936).
[3] R. Landshoff, Zeits. f. Physik **102**, 201 (1936).
[4] D. R. Hartree, Proc. Roy. Soc. **A141**, 282 (1933).

12. A Theory of Liquid Structure. JOSEPH HIRSCHFELDER, DAVID STEVENSON AND HENRY EYRING, *Princeton University*.—A simple model of the liquid is used to extend the equation of state previously obtained and to treat the process of fusion, viscous flow, and binary liquid systems. Our equation of state which applies to dense liquids has been fitted on to Happel's modification of van der Waals' equation to give a single equation applicable over the entire range from gas to liquid. A liquid differs from a solid in that the surplus volume in one part of the liquid becomes available in another part without an activation energy appreciable as compared to kT. This communal sharing of volume gives rise to an entropy of fusion R modified of course if there are other structural changes. Other entropy changes arise from expansion, changes of librations into free rotations and from polymerization. Double molecules held together by van der Waals forces are considered quantitatively and used in the explanation of viscous flow and deviations of the equation of state at the critical point. An explicit expression is given for the osmotic pressure of a binary liquid mixture. "Holes" are used to complete the analogy between critical phenomena for a one-component system and critical solution phenomena of binary liquids. In binary liquids the presence of a lower critical solution temperature above which two phases exist results from hydrogen (or analogous) bonds or bridges between unlike molecules which prevent free rotation. The critical mixing point coincides with the onset of free rotations which disrupt these bonds.

13. On the Magnetic Susceptibilities of Ti, V, and Cr Alums. J. H. VAN VLECK, *Harvard University*.—Lipson and Beevers' x-ray measurements on the alums seemingly indicate a field around the paramagnetic ion of trigonal, in fact nearly cubic symmetry.[1] With such a model, however, the excited states of Cr^{+++} cannot have the gyromagnetic ratio $g=2$ *uniformly* revealed by the Zeeman measurements of Spedding and Nutting.[2] Also it is then practically impossible to understand how the alums of Ti and V can *both* obey Curie's law with approximately the spin-only value of the susceptibility.[3] The dilemma is solved by using Jahn and Teller's neat observation that with orbital degeneracy the equilibrium position is "off-center," so that the symmetry is lowered and this degeneracy lifted.[4] Thus one can have a rhombic splitting of the order 10^3 cm^{-1} demanded by the susceptibility data and still retain the higher x-ray symmetry. The clusters about a given ion "resonate" through different rhombic arrangements and the average distribution revealed by x-rays can be cubic although instantaneously the field is only rhombic.

[1] Proc. Roy. Soc. **148A**, 664 (1935).
[2] Spedding and Nutting, J. Chem. Phys. **3**, 369 (1935).
[3] Cf. Siegert, Physica **3**, 85 (1936).
[4] Jahn and Teller, Phys. Rev. **49**, 874 (1936): Proc. Roy. Soc. in press.

14. Conductivity of Monovalent Metals. J. BARDEEN, *Harvard University*.—The conductivity of monovalent metals is computed from the Bloch theory under the assumption that the wave functions of the electrons are nearly the same as those of free electrons throughout the major part of the volume. The perturbation potential resulting from the Debye elastic waves, which produce transitions between the electronic states, is the sum of two terms: (1) the change in the total potential of the ions, which are assumed to move rigidly with the elastic waves, and (2) the change in the potential of the self-consistent field of the valence electrons. The second part tends to compensate the first. Scattering of the electrons through angles greater than $2\sin^{-1}2^{-1}$ results from the "Umklapp-

prozesse" of Peierls. It is shown that the probability of these transitions joins smoothly with the probability of transitions of the ordinary type, so that the probability is a continuous function of the angle of deflection, and transitions through all angles are possible. Comparison with experiment is made through Bethe's interaction constant C, which is a measure of the average scattering power of the elastic waves. The table gives the theoretical and experimental values of C/ζ, where ζ is the Fermi energy.

	Na	K	Cs	Rb	Cu	Ag	Au
C/ζ (theor.)	0.77	0.77	0.80	0.82	0.80	0.83	0.90
C/ζ (exp.)	0.77–1.04	0.87–1.04	1.3	1.6	1.32	1.21	1.30

15. The Gamma-Ray Spectrum of B^{10}. P. Gerald Kruger and G. K. Green, *University of Illinois.*—The gamma-rays from beryllium bombarded with deuterons have been analyzed by the examination of the Compton electrons ejected from a thin mica foil. A set of 1200 expansion chamber photographs yielded 480 acceptable tracks which were grouped and analyzed by the method of Richardson and Kurie.[1] The spectrum consists of six lines at 0.42, 1.15, 1.44, 2.17, 2.86 and 3.20 Mev±0.10 Mev. Bonner and Brubaker[2] have investigated the neutron spectrum of this reaction and found excited levels of the B^{10} nucleus at 0.55, 2.15 and 3.45 Mev±0.20 Mev. The set of gamma-ray lines represents all possible transitions between these levels and the ground state. Energy level values which best fit the gamma-ray spectrum are 0.5, 2.1 and 3.3 Mev, which is in good agreement with the results of Bonner and Brubaker.

[1] Richardson and Kurie, Phys. Rev. **50**, 999 (1936).
[2] Bonner and Brubaker, Phys. Rev. **50**, 308 (1936).

Paper No. 16 will be read by title.

16. Energy Loss of High Energy Beta-Rays in Lead. J. J. Turin and H. R. Crane, *University of Michigan.*—With the beta-rays from Li^8, we have continued our cloud chamber measurements[1] of the loss of energy of electrons in passing through lead, and have extended the data up to 11 Mev. A preliminary measurement of 415 tracks indicates the following losses in a ½ mm lead absorber.

Incident energy	Average loss
4 to 6 Mev	1.82 Mev
6 to 8 Mev	2.49 Mev
8 to 11 Mev	2.70 Mev

These values, like our values for the lower energies, are roughly 50 percent greater than predicted theoretically for ½ mm lead. Because of nuclear scattering in the lead the actual length of path of the electrons is expected to be appreciably greater than the thickness of the lead. A path length-thickness ratio of 1.5 might not be unreasonably large, and would bring the experimental values into fair agreement with theory.

[1] Abstract, Washington Meeting, 1937.

17. The Disintegration of Be by Protons and Deuterons. J. H. Williams, R. O. Haxby and W. G. Shepherd, *University of Minnesota.*—The yield curves for Be under proton and deuteron bombardment have been observed between 100 and 250 kv. Numbers *vs.* range curves for the various deuteron disintegration products are being investigated.

18. The Gamma-Rays from B and Be Under Proton Bombardment and from Li Under Deuteron Bombardment. W. G. Shepherd, R. O. Haxby and J. H. Williams, *University of Minnesota.*—Boron under proton bombardment emits a hard gamma-ray. The resonance observed by Bothe and Gentner at 180 kv has been verified. No gamma-rays harder than background x-rays from the accelerating tube have been observed from Be under bombardment with 200 kv protons. The gamma-ray previously reported from Li under deuteron bombardment has been found from absorption measurement in lead to have an energy of 400±25 kv. This energy agrees well with the observed difference in energy between the associated proton groups.

19. Delayed Alpha-Particles from Li^7 Bombarded by Deuterons. R. B. Roberts and L. R. Hafstad, *Department of Terrestrial Magnetism, Carnegie Institution of Washington,* and L. H. Rumbaugh, *Bartol Research Foundation of the Franklin Institute.*—The delayed alpha-particles emitted by Li^7 after bombardment with deuterons have been investigated, using a thin target (70 kv) deposited on a wheel which rotated continuously in the vacuum during bombardment. The alpha-particles counted were those emitted when the wheel had rotated 90° from the point of bombardment. In this manner it was possible to obtain numbers of delayed alpha-particles sufficient for accurate counting without any background of other heavy-charged particles. The excitation-curve of the delayed alpha-particles follows exactly that of the beta-particles in the voltage-range measured (550 to 950 kv) and shows the same break at 770 kv. The range-number curve of the delayed alpha-particles has been measured from 0.45 to 3.25 cm and shows a smooth, rapidly falling characteristic, with no evidence of any distinct groups of particles. If the logarithm of the number of alpha-particles observed at a given range is plotted against the corresponding energy, a straight-line relation holds over a factor of four in the energy.

20. The Range of Protons in Aluminum and in Air. D. B. Parkinson, R. G. Herb, J. C. Bellamy, C. M. Hudson, *University of Wisconsin.*—With the use of the two million volt generator developed at this laboratory the range of protons in air and in aluminum has been measured as a function of proton energy up to about 2 Mev. The range in air data are in good agreement with the theoretical values of Mano at high energies but diverge considerably at energies below 0.7 Mev. The ratio of the range in air to the range in aluminum has been found to increase from a value of approximately 1000 at 200 kv to a value of 1550 at 1200 kv. From 1200 kv up to the maximum voltage available the ratio is shown to remain nearly constant.

21. On Fluctuations in Size of Multiplicative Showers. W. H. Furry, *Harvard University.*—Recent calculations of Carlson and Oppenheimer and of Bhabha and Heitler

have made it evident that many features of cosmic-ray absorption and of shower production may be accounted for by assuming the validity of the existing radiation theory up to energies enormously higher than those at which a failure of the theory was formerly suspected. Being based essentially on calculations of the *mean numbers* of secondary photons and electrons produced under given conditions, these calculations need in some cases to be supplemented by a consideration of the fluctuations around these mean values. For example, such considerations are of importance in trying to determine to what extent the observed occasional production of large showers from fairly thin layers of lead ($\sim$1 cm) can be accounted for by the multiplication hypothesis. Bhabha and Heitler suggest, on apparently inadequate grounds, that the probability of appearance of n particles is given by the Poisson law, $P(n)=(m^n/n!)e^{-m}$, where m is the mean number. Detailed consideration of some drastically simplified but still presumably apposite cases has indicated that the Poisson law does not hold, and that the actual distribution is essentially of the type of $P(n)=m^{-1}(1-m^{-1})^{n-1}$. The latter law gives definitely greater probabilities for large showers.

22. The Hard Component of Cosmic Radiation. ARTHUR BRAMLEY, *Washington, D. C.*—The experiments indicate that the particle responsible for the hard component is extremely penetrating.[1] Since its ionization is nearly normal, the radiation losses must be abnormally low for an electron of that energy. The small cross section for radiation can arise either (1) because the particle has a greater mass than an electron, or because the electron has its charge distributed over a region with dimensions $\sim 10^{-11}$ cm, or (2) because the particle, if an electron, has a spin greater than $\frac{1}{2}$.[2] The radiation losses from an electron with spin greater than $\frac{1}{2}$ are considerably reduced in the case that its rest mass can be neglected in comparison with the mass associated with its kinetic energy. An extrapolation of the Fermi interaction to an electron of this character would lead us to expect that the maximum interaction would occur at an energy lower than $137\,mc^2$. It is impossible to predict what the strength of this interaction may be. It might turn out that the maximum strength of the interaction was such that all the low energy electrons with spin greater than $\frac{1}{2}$ would be absorbed into the nuclei of the medium they were traversing.

[1] Street and Stevenson, Abstract 40, Bulletin, Vol. 12, No. 2: Anderson, Science News, November, 1936.
[2] Dirac, Proc. Roy. Soc. July (1936).

23. The Scattering of Neutrons by Iron. R. F. BACHER, *Cornell University.*—The scattering of rhodium resonance neutrons by Fe has been studied by using thin and thick iron scatterers above a detector on a paraffin cylinder. It has been found that the albedo effect contributes greatly to the scattering. In order to take this into account small thick detectors were used to determine the number of neutrons passing in each direction through the interface. For 5.4 cm Fe the fraction of neutrons reflected was 0.40 while for 0.6 cm Fe the fraction reflected was 0.20. For the thick iron this reflection gives a ratio of absorption to scattering cross section $\sigma_a/\sigma_s=0.22$. From the measurements on the thin plate $\sigma_s=11\times10^{-24}$ cm^2.

24. On the Scattering of Neutrons by Magnetic Dipoles. OTTO HALPERN AND M. H. JOHNSON, *New York University.*—In establishing the theoretical basis for experiments previously suggested by the authors (to appear in the *Physical Review*) we have made a more detailed investigation of the scattering of neutrons by magnetic dipoles. In these calculations the atomic dipoles were treated quantum-mechanically leading to significant differences from previously suggested formulae. In general, in paramagnetic as well as ferromagnetic substances the magnetic scattering cross sections are increased because of collisions of the second kind whereby transitions of the atom are induced either among the different magnetic sublevels or to neighboring spectroscopic levels. Polarization of the neutron beam may occur not only because of interference between nuclear and magnetic scattering, but also because of the magnetic scattering alone if transitions among the magnetic sublevels are possible. The latter type may assist or diminish the former depending upon the phase change produced in the nuclear scattering (real or virtual level of the system nucleus plus neutron). Other changes in the polarization effects can be produced by spin dependent forces and the presence of isotopes. The additional transitions here discussed tend to diminish polarization effects in ferromagnetic bodies.

25. Simple Derivation of the Formula for the Total Scattering of X-Rays from a General Crystal. G. G. HARVEY, *Massachusetts Institute of Technology.*—For the case of a general triclinic crystal having any number of atoms in the unit cell there is derived a formula for the intensity of total scattering of x-rays in any direction. In those directions satisfying the Laue equations the formula reduces to the usual one while in other directions the formula gives the intensity of the diffuse scattering. The derivation is purely classical; each electron in the various atoms is assumed to scatter according to the J. J. Thomson formula and the amplitude of the resultant scattered wave is the vector sum of the amplitudes scattered by the individual electrons, the sum being taken over the entire crystal. The observed intensity is obtained by averaging over the orbital motion of the electrons and the thermal vibration of the atoms. No assumption is made as to the nature of either of these motions. This procedure leads to the *total intensity* (coherent+incoherent scattering) and from the nature of the resulting equation the incoherent part may be identified. As is well known, such a derivation by classical methods leads to the same result as a quantum-mechanical treatment except for corrections due to relativity, etc. Since these occur only in the incoherent part of the scattering they may be easily added.

26. The Reflection Coefficient of Mercury. R. C. MASON, *Westinghouse Electric & Manufacturing Company.*—Two new methods of measuring the rate of evaporation of liquid mercury show that the rate may be small compared to the kinetic theory calculated maximum rate, even

though the mercury surface may appear clean and bright to the eye; this result may be interpreted as the mercury atoms having a large reflection coefficient. The first method uses a modified Chattock micromanometer to measure the change in pressure on a mercury pool when it is suddenly exposed to a nearby liquid air cooled condensing surface, and also to indicate the subsequent evaporation rate. From the two, the reflection coefficient may be calculated, independent of the calibration of the instrument, the temperature of the mercury surface, or the vapor pressure of mercury. Values of reflection coefficient from 0.93 to over 0.99 were obtained in this way. The second method employs an ionization gauge to measure the density of vapor between the evaporating mercury pool and the condensing surface. When the temperature of the mercury is measured by a thermocouple, the rate of evaporation may be found. Corresponding reflection coefficients from 0 to almost 1 have been observed.

27. The Infrared Absorption Spectra and Approximate Force Constants of Propene and Allene. LYMAN G. BONNER* AND ROBERT HOFSTADTER, *Princeton University.*—The present report marks the beginning of a systematic survey of the C=C force constant as it appears in various organic compounds. The infrared absorption spectra of gaseous propene and allene have been obtained with a rocksalt prism spectrometer in the region 1–15.5μ. Runs were taken with a 15 cm cell and under various pressure conditions ranging from 35 to 700 mm. Our measurements show bands in propene at 6211, 4386, 3021, 2309, 2008, 1815, 1634, 1435, 1183, 984, and 906 cm^{-1} and in allene at 6289, 4525, 3086, 2392, 1961, 1678, 1381, 1143, 1033, and 840 cm^{-1}. Approximate force constants have been calculated by the method of rigid groups, treating propene as an asymmetric triatomic molecule and allene as a symmetric linear-triatomic. The results of these calculations give the double bond force constant as 10.0 and 9.45×10^5 dynes/cm, respectively, in the two cases. For propene the calculations also furnish a value for the C–C force constant, in this case 3.8 in the above units.

* National Research Fellow.

28. Index of Refraction of Methane in the Infrared. R. ROLLEFSON AND R. J. HAVENS, *University of Wisconsin.*—The apparatus which was built to measure the index of refraction of HCl in the near infrared[1] has been improved so that it is now possible to make measurements out to 15μ. This improved apparatus has been used to measure the index of refraction of methane from 1μ to 15μ. The variation of the index in the neighborhood of the 3.3μ and 7.7μ bands shows that the ratio of the intensities of these bands is approximately 2.3. The overtone bands in the vicinity of 2.3μ contribute a noticeable but small amount to the index of refraction. No quantitative estimate of their intensity can be made, however, since several overtone bands occur in a narrow region of the spectrum where the rock salt monochromator has its smallest dispersion.

[1] Phys. Rev. **48**, 779 (1935).

29. The Products of Dissociation of Ethylene by Electron Impact. JOHN T. TATE, A. HUSTRULID AND P. KUSCH, *University of Minnesota.*—The dissociation of ethylene by electron impact has been investigated with a mass spectrograph. The ions which have been identified, their relative intensities at an electron energy of 75 volts and their appearance potentials are tabulated. It was not possible

Ion	Int.	A.P. (volts)	Ion	Int.	A.P. (volts)
$C_2H_4^+$	100	10.80±0.05	CH^+	2.5	22.9±0.5
$C_2H_3^+$	57.2	14.1 ±0.1	C^+	1.7	24.6±0.5
$C_2H_2^+$	51.3	13.4 ±0.2	H_2^+	0.13	22.4±1.5
C_2H^+	8.8	19.2 ±1.0	H^+	0.5	26.2±1.5
C_2^+	2.7	26.4 ±1.0			28.2±1.5
CH_4^+	0.06	——	$C_2H_4^{++}$	0.3	——
CH_3^+	0.6	——	$C_2H_3^{++}$	0.3	36.0±1.0
CH_2^+	5.2	19.2 ±0.3	C_2H^{++}	0.006	——

to identify the doubly ionized ions in every case, nor to determine their appearance potentials because of coincidences between these peaks and those due to singly ionized ions. The peaks CH_4^+ and H_2^+ have satellites whose appearance potentials are higher than those due to the main peak. These peaks may be accounted for by the assumption that two processes are responsible for the formation of these ions, with one of which a considerable amount of kinetic energy is associated.

30. The First Spark Spectrum of Tin: Sn II. W. W. MCCORMICK AND R. A. SAWYER, *University of Michigan.*—The low vapor pressure of tin makes the excitation of the higher terms of its spark spectrum difficult. Earlier investigators, notably Lang[1] and Green and Loring,[2] have been able to establish with certainty only a few of the lowest terms, with not more than two terms in any series. Thus, even the value of the ionization potential has been only approximate. The spectrum has been remeasured throughout the whole photographic region from 800A to 10,000A by using a hollow cathode discharge and long exposures. It has been possible to extend all the known series. In addition, the *F–G* series has been discovered, and five members found. The close agreement of the *G* terms with the Rydberg formula permits the determination of the ionization potential of Sn II as approximately 118,010 cm^{-1}, a change of about 300 cm^{-1} from the previous estimate. In all, about 80 new lines have been classified, determining about 25 new terms.

[1] Phys. Rev. **35**, 445 (1930).
[2] Phys. Rev. **30**, 574 (1927).

31. Spectral Series in Neutral Tungsten, W I, and Limitations Upon the Possibility of Determining Spectroscopic Ionization Potentials. O. LAPORTE, *University of Michigan,* AND J. E. MACK, *University of Wisconsin.*—The identification of $5d^46s7s\ ^7D_{12345}\ ^5D_{04}$ and $5d^46s8s\ ^7D_4$? leads to the series limit value W I $5d^46s^2\ ^5D_0$ (normal state) $-$ W II $5d^46s\ ^6D_{\frac{1}{2}}=7.84$ volts; but probably the lowest level of W II is $5d^5\ ^6S$, some tenths of a volt below $^6D_{\frac{1}{2}}$, leading to a tentative value of 7.6±0.2 volts for the ionization potential. Among the complex spectra of atoms heavier than silver, series have been found only in Sa, Eu, and W. It seems quite likely that they can never be

identified among the configurations involving partly filled d-shells ($n>4$) or f-shells except near the ends of the periods and near the very middle, i.e., with $d^5(^6S)nl$, $f^7(^8S)nl$, $d^4\,^5D$, $d^6\,^5D$, $f^6\,^7F$, or $f^8\,^7F$, alone or with one or two added s-electrons.

32. Collisions of the Second Kind and Sensitized Fluorescence of Tin and Mercury Atoms. J. G. WINANS AND R. M. WILLIAMS, *University of Wisconsin.*—Several tin lines have been observed in a low voltage Tesla discharge and in sensitized fluorescence. The discharge was produced in a fused quartz tube containing mercury and tin, and heated by a Bunsen burner. Tin lines were emitted from a region locally heated to about 900°C with a blow torch. In comparison with the arc spectrum of Sn, the following Sn lines in this discharge were enhanced over neighboring lines: 3262, 3034, 2840, and 2430. All enhanced lines except 3034 originate from Sn levels nearly equal in energy to Hg levels. Sensitized fluorescence of Sn was obtained by heating a drop of tin in the tube with a hydrogen-air blow torch and illuminating with light from a Hanovia SC-2537 tube. Mercury pressure was regulated with a Bunsen burner. The Sn lines observed are listed below with estimated intensities in parentheses. 4524 (2), 3801 (3), 3331 (4), 3262 (10), 3175 (4), 3034 (10), 3009 (2), 2863 (3), 2840 (10), 2706 (10), 2661 (4), 2547 (2), 2430 (4). All strong Sn lines excepting again 3034 arise from levels with low energy discrepancy with Hg 2^3P_2, 2^3P_1, or 2^3P_0. The average thermal energy available for excitation was 0.101 volt. The pressure of Sn vapor was less than 5.6×10^{-5} mm of mercury.

33. The Perpendicular Vibration Bands of NH_3. E. F. BARKER, *University of Michigan.*—The two parallel vibrations of NH_3 are well known both from infrared and from Raman observations. They yield bands at 3μ (ν_1) and at 10μ (ν_3). The intense band at 6μ corresponds to one of the perpendicular vibrations (ν_4). Its structure has been observed in detail, and is found to correspond exactly with the theoretically predicted pattern, both as to locations and as to intensities of the lines. The center is at 1626 cm^{-1}. The position of the other fundamental (ν_2) has not previously been determined. By means of two combination bands, one at 2.2μ ($\nu_2+\nu_3$) and the other at 4.0μ ($\nu_2-\nu_3$) the characteristic structure of which can be recognized, it is now possible to determine this band exactly. Its center is at 3407 cm^{-1}. The fundamental itself is weak, and is difficult to observe because of the intense parallel band ν_1 which overlies it. A band has also been found corresponding to the Raman line observed in gaseous ammonia at 3219 cm^{-1}. This is interpreted as the parallel component of $2\nu_4$.

34. X-Ray Spectroscopic Data in Regard to the Electronic Energy Bands in Potassium and Sodium Chloride. JOSEPH VALASEK, *University of Minnesota.*—The K beta lines of chlorine in the chlorides of sodium and potassium have been measured. Long exposures were given to record the weaker lines and an attempt has been made to fit the group into an energy band scheme of the usual type. The principal line offers little difficulty and gives the binding energy of the Cl^- $3p$ electrons. The longer wave-length satellites may come from the Na^+ $2p$ and the Cl^- $3s$ levels, though the numerical agreement with published calculations is not particularly good. The first two satellites on the short wave-length side indicate energy levels at elevations above Cl^- $3p$ which agree roughly with the wave-lengths of the F and U bands in the colored crystal. The other short wave-length satellites are probably due to double ionization.

35. The Second Order Null Result of the Michelson-Morley-Miller Experiment. W. B. CARTMEL, *Université de Montréal.*—In Lorentz's triangle the two rays from a point P meet again at the same point P on the plate. In Silberstein's triangle[1] the two rays from the point P meet again at a point P' in the ether, at a distance lv^2/c^2 beyond the second position of P, and there is therefore no v^2/c^2 difference of phase between the two wave fronts that intersect at the point P'. This justifies Lorentz's theorem. The wave front reflected at the point where ray BP' meets the plate is parallel to the wave front in arm A that passes through P', but there is a third order difference in position. The third order term is due to the use of a fixed source instead of a terrestrial source. The same v^3/c^3 effect comes in if we use an inclined ray from a terrestrial source, in which case the experiment is not ideal. A fringe shift is indicated by the ray BP' in Silberstein's triangle meeting the plate at a distance lv^2/c^2 from the point P', but this is almost annulled. Lengthening arm B would cause the ray to move further up the plate and increase the fringe shift. Neither arm is actually lengthened but there is an increase in path length in arm A relatively to arm B, and this moves the end point of the ray *down* the plate so that the rays PBP and PAP meet at the plate as in Lorentz's triangle. The $\frac{1}{2}v^2/c^2$ term in Silberstein's equation comes from the virtual rotation of the 45° plate which moves the ray up the plate, and the virtual change in length of arm brings it back again, giving us a v^2/c^2 null result, together with a v^3/c^3 fringe shift.

[1] L. Silberstein, *The Theory of Relativity* (Macmillan, London, 1914), p. 89.
[2] H. A. Lorentz, *The Theory of Electrons* (Stechert, New York, 1916), p. 181.

36. Is the Null Result Inherent in the Interferometer or in the Motion? W. B. CARTMEL, *Université de Montréal.*—Let us consider the motion of the observer and of the source separately, as was done by Voigt.[1] We shall take the motion of the observer to be along a tangent to a great circle on one of the concentric spherical wave surfaces emanating from a point source S. The ray at right angles to the motion meets the observer at the point where the tangent meets the radius, the observer's motion beginning at the point of intersection of the secant and the tangent. The source, moving with the same speed as the observer, finds itself at a point S', at the end of the longer diagonal of a parallelogram, at the instant of time when the ray meets the observer. The two sides of the parallelogram give us the two positions of the apparent

ray joining source and observer. Measuring from S' along the apparent ray to the wave surface, we shall find a distance $l(1+v^2/c^2+\frac{1}{4}v^4/c^4+\cdots)$, and from the same point, along the apparent ray in the direction of the motion, a distance $l(1+v^2/c^2+0v^4/c^4+\cdots)$. To the second order our wave surfaces are circles, a result similar to Voigt's.[1] This is on axes fixed in the ether. On axes moving with the earth the radius to the wave surfaces is given by the latter expression in both directions, which is identical with a result found with the interferometer,[2] but inherent in the motion.

[1] W. Voigt, Göttinger Nachrichten, 1-21, 41 (1887).
[2] W. B. Cartmel, Phys. Rev. **49**, 647, 649 (1936).

Author Index to Abstracts of Madison Meeting

Bacher, R. F.—No. 23
Baldes, E. J.—see Herrick, J. F.
Bardeen, J.—No. 14
Barker, E. F.—No. 33
Bellamy, J. C.—see Parkinson, D. B.
Bonner, Lyman G. and Robert Hofstadter—No. 27
Bramley, Arthur—No. 22

Cartmel, W. B.—No. 35
—— No. 36
Cashman, R. J.—No. 10
Cole, K. S.—see Curtis, H. J.
Crane, H. R.—see Turin, J. J.
Curtis, H. J. and K. S. Cole—No. 3

Depp, Wallace A.—see Reich, Herbert J.

Eyring, Henry—see Hirschfelder, Joseph

Furry, W. H.—No. 21

Green, G. K.—see Kruger, P. Gerald
Greenfield, E. W.—No. 6

Hafstad, L. R.—see Roberts, R. B.
Halpern, Otto and M. H. Johnson—No. 24
Harvey, G. G.—No. 25
Havens, R. J.—see Rollefson, R.
Haxby, R. O.—see Shepherd, W. G.
—— see Williams, J. H.
Herb, R. G.—see Parkinson, D. B.
Herrick, J. F. and E. J. Baldes—No. 4
Hirschfelder, Joseph, David Stevenson and Henry Eyring—No. 12
Hofstadter, Robert—see Bonner, Lyman G.
Hudson, C. M.—see Parkinson, D. B.
Hustrulid, A.—see Tate, John T.

Johnson, M. H.—see Halpern, Otto

Kruger, P. Gerald and G. K. Green—No. 15
Kusch, P.—see Tate, John T.

Landshoff, Rolf—No. 11
Laporte, O. and J. E. Mack—No. 31

Mack, J. E.—see LaPorte, O.
Mason, R. C.—No. 26
McCormick, W. W. and R. A. Sawyer—No. 30

Parkinson, D. B., R. G. Herb, J. C. Bellamy and C. M. Hudson—No. 20
Pielemeier, W. H.—No. 1
—— No. 2

Reich, J. and Wallace A. Depp—No. 7
Roberts, R. B., L. R. Hafstad and L. H. Rumbaugh—No. 19
Rollefson, R. and R. J. Havens—No. 28
Rumbaugh, L. H.—see Roberts, R. B.

Sawyer, R. A.—see McCormick, W. W.
Shepherd, W. G.—see Williams, J. H.
——, **R. O. Haxby and J. H. Williams**—No. 18
Skinner, Selby M.—No. 5
Stevenson, David—see Hirschfelder, Joseph
Suits, C. G.—No. 8

Tate, John T., A. Hustrulid and P. Kusch—No. 29
Theim, Louis P., S.M.—No. 9
Turin, J. J. and H. R. Crane—No. 16

Valasek, Joseph—No. 34
Van Vleck, J. H.—No. 13

Williams, J. H.—see Shepherd, W. G.
——, **R. O. Haxby and W. G. Shepherd**—No. 17
Williams, R. M.—see Winans, J. G.
Winans, J. G. and R. M. Williams—No. 32

Proceedings of the American Physical Society

MINUTES OF THE DENVER MEETING, JUNE 25 AND 26, 1937

THE 215th regular meeting of the American Physical Society was held in the Y. W. C. A. auditorium at Denver, Colorado on Friday and Saturday, June 25 and 26, 1937.

Morning sessions were devoted to symposia, and the afternoons to the presentation of the twenty-six contributed papers. Papers 5, 6, 7, 11, 17, 22 and 26 were read by title.

The symposium of Friday morning, held jointly with the Astronomical Society of the Pacific and the Section on Astronomy (D) of the A.A.A.S., was concerned with *Astrophysical Problems of the Ionosphere*. Speakers were:

N. E. BRADBURY, *Stanford University*
A. G. MCNISH, *Carnegie Institution of Washington*
J. H. DELLINGER, *National Bureau of Standards*
R. S. RICHARDSON, *Mount Wilson Observatory*

On Saturday morning a symposium on *Cosmic Physics and High Altitude Effects* was conducted as a joint session with the Section on Physics (B) of the A.A.A.S. The speakers were:

JOYCE C. STEARNS, *University of Denver*
THOMAS H. JOHNSON, *Bartol Research Foundation*
HURD C. WILLETT, *Massachusetts Institute of Technology*
J. C. STREET, *Harvard University*

Maximum attendance at the sessions ranged from 60 to 75 persons.

Abstracts of the contributed papers are given in the following pages. An Author Index will be found at the end.

PAUL KIRKPATRICK, *Local Secretary for the Pacific Coast*

ABSTRACTS

1. High-Pressure Ionization Currents at High and Low Temperatures. JAMES W. BROXON AND GEORGE T. MERIDETH, *University of Colorado.*—The investigation of weak gamma-ray ionization currents in air, discussed in an earlier paper,[1] has been extended to nearly 200 atmospheres. Measurements have also been made in the neighborhood of 96°C and −76°C. Under no conditions have saturation currents been obtained at the highest gradients employed. At room temperature, the ionization-pressure curves appear to pass through maxima within the new pressure range. At 96°C the high-gradient ionization currents continue to increase with pressure at the highest pressures. At −76°C, the ionization-pressure curves have decided maxima in the neighborhood of 60 atmospheres. At the new temperatures the negative ionization currents continue to exceed the positive under the same conditions.

[1] James W. Broxon and George T. Merideth, Phys. Rev. **49**, 415 (1936).

2. Preliminary Report on Recombination of Ions in Air at High Pressures. GEORGE T. MERIDETH AND JAMES W. BROXON, *University of Colorado.*—In conjunction with the measurements of ionization currents induced by weak gamma-rays in air at high pressures (see preceding abstract), determinations of the charge density of ions of one sign have been made also. Analysis of the growth curves obtained from these data by means of the customary equation, $dn/dt = q - \alpha n^2$, and also the equation suggested by the experience of Power,[1] $dn/dt = q - \beta n$, has been begun.

At present it appears that the data conform more closely to the latter equation, although there are indications that the equation proposed by Schweidler[2] would represent the results still more closely.

The values of the ordinary recombination coefficient, α, computed by a method of approximations are in fair agreement with those determined by Lea[3] by a somewhat similar method.

[1] A. D. Power, J. Frank. Inst. **196**, 327 (1923).
[2] Schweidler, Akad. Wiss. Wien. Ber. **127** (2a), 953 (1918).
[3] D. E. Lea. Proc. Camb. Phil. Soc. **30**, 80 (1933).

3. The Recombination of Ions in Pure Oxygen as a Function of Pressure and Temperature. M. E. GARDNER, *University of California.* (*Introduced by Leonard B. Loeb.*)—The coefficient of recombination, α, for the pure oxygen (no free electrons) showed α to be a constant independent of the age of the ion from 10^{-2} sec. to 1 sec. The value of α

for pure oxygen at room temperature and atmospheric pressure was $(2.0\pm0.1)\times10^{-6}$. The value of α as a function of pressure and temperature was in agreement with the theoretical values of J. J. Thomson (molecular weight of ions 64, mean free path $\frac{1}{8}$ that for molecules) and completely incompatible with the theory of Langevin. The pressure variation was in good agreement from 10 cm to 76 cm pressure. Below 10 cm the high diffusion rate and the presence of free electrons vitiated the results obtained. The temperature variation was in good agreement from -40°C to 100°C. At -78°C the experimental values deviated from the theoretical value but could be made to agree if it were assumed that the molecular weight of the ions increased to 96. These experiments show that initial and preferential recombination is occurring to some extent even at 0.1 second. This indicates diffusion coefficient for the ions from a paired configuration of the order of 100 times the normal coefficient. A similar conclusion was drawn by Loeb.[1]

[1] Loeb, Phys. Rev. **51**, June 15 (1937).

4. Ionization in Cosmic Ray Tracks. MERLE A. STARR, *University of California.* (*Introduced by Robert B. Brode.*)—According to the theory of Bethe and Heitler, the ionization produced by electrons should be a function of energy with a minimum at 10^6 ev and having a value at 10^5 and 10^6 ev twice as great as the minimum. An attempt has been made to observe such a function by counting drops in diffuse tracks about one second old, photographed in a counter controlled cloud chamber. Measurements show that on the average, particles of near 5×10^3 Hρ do not ionize as heavily as those of lower and of higher energies, but tracks above the limit of energy measurement, 2×10^6 Hρ, show a wide fluctuation of from 20 to 60 drops per cm path. A few straight tracks with approximately twice the average ionization have been observed in small showers. These may be produced by very high energy electrons or by particles of greater mass. Since the low drop count has not been traced to faulty expansion of the cloud chamber, it appears likely that such tracks were produced by electrons, and the most dense by protons of high energy. A more detailed study of recombination of ions will make it possible to determine whether the effect is real.

5. The Townsend Coefficient in Pure Nitrogen. WOOLFORD E. BOWLS, *University of California.* (*Introduced by L. B. Loeb.*)—An investigation of the effect of cathode material on β, the second Townsend coefficient of ionization, has led to surprising results with Pt and Na cathodes. The values of α, the first Townsend coefficient of ionization, obtained in an outgassed, mercury free chamber were 20 percent less and 5 percent greater for Pt and Na cathodes, respectively, than values obtained by Townsend in the presence of Hg vapor. For the Pt cathode, and $X/p<400$, the values of β/p were larger than those found by previous investigators using mercury coated surfaces. The values of β/p rose sharply from zero at $X/p=65$ until a plateau was reached extending from $X/p=140$ to $X/p=240$ where $\beta=0.05$. The curve then rose into the Townsend curve at about $X/p=400$. For the Na cathode, the β/p curve lay entirely below the Pt curve, starting at an $X/p=140$ and rising slowly to a less pronounced plateau at an $X/p\sim325$. For higher X/p the curve essentially parallels the Pt curve. Plots of β/α as a function of X/p showed a sharp peak in each curve, possibly due to photoelectric phenomena in the low X/p range, followed by a gradual rise, resulting from secondary emission due to positive ion bombardment.

6. The Volta Potential between Barium and Magnesium. An Experimental Test of the Relation between Work Functions and Volta Potential. PAUL A. ANDERSON, *State College of Washington.*—Recent photoelectric determinations of the work functions of barium[1] and of magnesium[2] in which the technique of preparation of the metal surfaces has been similar to that used by the author in measurements of contact differences of potential[3] have suggested the possibility of checking the Sommerfeld-Eckart relation under unusually favorable experimental conditions. A first measurement on fractionally and multiply distilled surfaces of Ba and Mg deposited on glass at room temperature has yielded the value 1.08 v for the contact difference of potential between these metals. The difference of 1.08 ev between the work functions predicted by the Sommerfeld-Eckart relation agrees to within 0.01 ev with the results of Jamison and Cashman, $\phi_{Ba}=2.51$–2.52 ev, and Cashman and Huxford, $\phi_{Mg}=3.60$ ev. While the exactness of this agreement is, no doubt, fortuitous to some degree it indicates that the results of independent photoelectric and contact potential measurements which are obtained under adequate experimental conditions may be used interchangeably in work function calculations.

[1] Jamison and Cashman, Phys. Rev. **50**, 624 (1936).
[2] Cashman and Huxford, Phys. Rev. **48**, 734 (1935); Mann and DuBridge, Phys. Rev. **51**, 120 (1937).
[3] P. A. Anderson, Phys. Rev. **47**, 958 (1935); 320 (1936).

7. Some Applications of Spectroscopy to Agriculture. STANLEY S. BALLARD, *University of Hawaii.*—A satisfactory technic for the qualitative analysis of various agricultural specimens has been developed. Although developed primarily for the analysis of soils, fertilizers and plant materials, the technic doubtless has a much wider field of application. Samples to be analyzed are burned in a carbon or copper arc, and their spectra are produced and photographed by a large quartz spectrograph. The photographic plates may be analyzed rapidly to establish the presence or absence in the sample of some fifty of the elements. Features of the technic concern the care exercised in the preparation of electrodes and samples to prevent the inadvertent introduction of any impurities. The procedure developed is at present being applied with success to the analysis of various commercial fertilizers for the presence of the rarer metallic elements, the so-called less essential plant nutrients. This program is being carried on in the spectroscopic laboratory of the Experiment Station of the Hawaiian Sugar Planters' Association with the assistance of P. E. Chu and P. L. Gow, Assistant Chemists at the Experiment Station. An article describing the experi-

mental procedure and some of its applications is being prepared for publication in an appropriate journal.

8. The Design and Performance of Quantity Heads for the Air-Driven Ultracentrifuge. RALPH W. G. WYCKOFF AND J. B. LAGSDIN, *Rockefeller Institute for Medical Research, Princeton, N. J.*—The ultracentrifugal sedimentation of antibodies and the similar isolation of virus proteins have demonstrated that many biologically important substances too unstable for chemical purification can be prepared by physical procedures based on the ultracentrifuge. The need that has accordingly arisen for handling comparatively large volumes of liquid has led to the design of rotors holding 100 to 200 cc. Studies have been made to determine the maximum speeds at which such rotors of several shapes and materials can be operated and to ascertain the variation with tube angle of the efficiency of sedimentation in ultracentrifugal fields. The information thus gained has been utilized in the construction of light metal alloy heads that can be employed for quantity centrifugation at fields up to 250,000 to 300,000 times gravity. These will be described and a discussion will be given of their performance when operated by the current design of air-turbine.

9. The Theory of Optical Activity. E. U. CONDON, WILLIAM ALTAR AND HENRY EYRING, *Princeton University.*—It is shown that a single electron moving quantum-mechanically in a field of suitable dissymmetry suffices to make a medium containing such molecules show optical activity. This is in striking contrast to the models studied by Born, Kuhn and others in which the rotatory dispersion is related to the hypothetical existence of coupled oscillators in the molecule. Here the absorption bands are associated with one-electron jumps in the molecule, the rotatory power arising from the fact that the electrons move in a dissymmetric field.

It turns out that to account for the presence of such a field one merely has to consider the presence of electrostatic charges in the vicinity of the optically active electron, in accordance with standard dipole data. This is confirmed by calculations of the rotatory dispersion in specific examples, using the following expression for the potential energy function:

$$V=\tfrac{1}{2}(k_1x^2+k_2y^2+k_3z^2)+Axyz,$$

where the constants k_1, k_2, k_3 and A may be calculated from the dipole moments. The numerical agreement with the experimental curves is very satisfactory.

10. Luminescence and Color Excited by Radium in Zinc Borate Glasses Which Contain Manganese. BYRON E. COHN, *University of Denver,* AND S. C. LIND, *University of Minnesota.*—A series of synthetic zinc borate glasses were made up to contain known concentrations of manganese between zero and five percent. These glasses were exposed to radium.

The thermoluminescence of the samples was determined. Manganese was found to act as an activator; the optimum concentration was at 0.05 percent manganese. The luminescence emission was in the form of a band with a maximum around 6000A.

The change in absorption of light due to the exposure of such glasses to radium was determined quantitatively for both the visible and ultraviolet regions. The violet color developed in these glasses by exposure to radium was found to be due to a preferential increase in the absorption for wave-lengths longer than 5000A. Exposure to radium produced an increased absorption in the ultraviolet region. The change in absorption depended both upon the wavelength and the concentration of manganese. At 2537A the greatest change in absorption was at 0.05 percent manganese and appeared related to the thermoluminescence emission.

11. The Structure of the Arc Spectrum of Tungsten, WI. J. E. MACK, *University of Wisconsin,* AND OTTO LAPORTE, *University of Michigan.*—The analysis of the spectrum of WI, begun by one of us,[1] has been brought to a certain stage of completion by the classification of all stronger and of many of the weaker lines into a diagram of 42 even and 195 odd levels. Although for the latter no quantum number other than J can be and should be assigned, this is possible in numerous cases for the even levels under either one of the following conditions: (a) An extreme J value identifies a level independently of coupling; (b) disturbing terms are absent. A check is obtained on the completeness of the level diagram found in this work by counting the levels of a definite J value and comparing these numbers with the theoretically expected ones.

[1] O. Laporte, Naturwiss. **13**, 627 (1925).

12. The Preparation of Auroral Afterglow Tubes. JOSEPH KAPLAN, *University of California at Los Angeles.*—The failure of Cario and Stille to reproduce some of the results obtained in our auroral afterglow experiments has led to a repetition of some of our former studies, and in particular to a more detailed study of the preparation of the tubes which show these glows. A series of spectrograms is shown which shows the cleanup of impurities during the preparation of several tubes. These afterglow spectra illustrate clearly three distinct, successive stages of the afterglow as the N_2 becomes purer, namely, the cyanogen stage; the cyanogen-nitrogen-molecule-ion stage; and the afterglow in pure nitrogen, which may be called the auroral stage. Following the auroral stage the tube cleans up the nitrogen itself, and a series of spectrograms of the afterglow is obtained which illustrates the remarkable variation in the auroral afterglow with pressure. In the cyanogen stage under certain conditions the afterglow and discharge are so distinct that an objective proof of the accuracy of our technique for photographing the afterglow has been obtained.

13. The Frequency of Cosmic-Ray Showers Produced in Different Metals. J. C. STEARNS, *University of Denver,* AND D. K. FROMAN, *MacDonald College, McGill University.*—A study of the frequency of shower production in lead, tin, copper, iron, zinc, and aluminium is being made. Three Geiger counters arranged in the form of a triangle and connected to a circuit which records triple coincidences are

being used to detect showers originating in blocks equal in mass and cross-sectional area. The counter voltage and temperature is kept constant. The number of coincidences was taken with shower producing blocks used above the counters in the following order: Al, air, X, X, air, Al, etc. X is any element except Al. The ratio (X-air) to (Al-air) was determined and from this the showers per atom (N) for each element determined. N varies approximately as the square of the atomic number.

14. The Vertical Cosmic-Ray Intensity up to 43.5 mm Hg. THOMAS H. JOHNSON, *Bartol Research Foundation of the Franklin Institute.*—A balloon flight on May 21 carried a double coincidence counter apparatus with radio transmitter to a height corresponding to barometric pressure 43.5 mm and down again. The radio signals reporting coincidences, barometric height and temperature were distinct throughout the entire flight to the ceiling and down to within a mile of sea level, a total time of two hours and forty-one minutes. Signals were received from a fire lookout tower of the Pennsylvania Department of Forests where interference noise is at a low level. Records of the signals were made automatically on photographic film and manually, from the audible sounds, on a telegraphic register. Details of the results will be ready at the time of presentation.

15. The Effects of Time and Meteorological Factors on the Intensities of Cosmic-Ray Primaries and Showers. DAROL K. FROMAN, *MacDonald College, McGill University,* AND J. C. STEARNS, *University of Denver.*—The hourly counting rate of a set of triple-coincidence Geiger-Müller counters has been observed for about 3000 hours in counting about 190,000 showers from lead, and for about 1500 hours in counting about 160,000 vertical primary rays. The counting rates, Y, have been fitted separately by the least-square method, to an equation of the form

$$Y=A_1\sin(2\pi/24)t_1+A_2\cos(2\pi/24)t_1+A_3\sin(2\pi/24)t_2 +A_4\cos(2\pi/24)t_2+A_5p+A_6T+A_7w+A_8H+A_9n,$$

where the symbols have the following significance: t_1, local solar hour; t_2, local sidereal hour; p, atmospheric pressure in cm of Hg; T, atmospheric temperature in °C; w, atmospheric humidity in g per cu. meter; H, horizontal component of the earth's magnetic field in gammas; n, serial number of the observation. The following table gives the partial regression coefficients, expressed as a percentage of mean counting rate, and their probable errors. The values marked with an asterisk differ significantly from zero.

	FOR SHOWERS	FOR PRIMARIES
A_1	$+3.8 \pm 2.5$	$+3.1 \pm 3.0$
A_2	$+0.2 \pm 2.5$	$+2.1 \pm 2.4$
A_3	-3.2 ± 2.7	-5.1 ± 3.1
A_4	-1.3 ± 2.2	-3.9 ± 2.3
A_5	$*-5.45 \pm 0.44$	$*-1.73 \pm 0.56$
A_6	$*+0.83 \pm 0.10$	$*+0.29 \pm 0.088$
A_7	-0.068 ± 0.13	$*-0.60 \pm 0.12$
A_8	$+0.018 \pm 0.011$	$+0.0049 \pm 0.014$
A_9	$*-0.033 \pm 0.0025$	$+0.0014 \pm 0.0030$

16. Alpha-Particle Counting and Geologic Ages. ROBLEY D. EVANS AND CLARK GOODMAN, *Massachusetts Institute of Technology.*—The complete theory of alpha-particle counting for thin and thick sources has been developed and experimentally confirmed by the method of internal standardization.[1] The uncertainty in the correction for internal absorption of alpha-rays within the source is large in the case of thick sources but becomes negligible for a thin source having a thickness of the order of one air-cm. For thin sources the alpha-ray counting rate becomes a direct measure of the rate of production of helium. Representative portions of geologic samples, ground to an average particle size of only a few microns, are temporarily suspended in ethyl alcohol and evaporated in an atmosphere of radioactive-free nitrogen onto a previously weighed aluminium disc. The alpha-ray emission of the resulting uniform compact deposit is then determined with a photographic recording vacuum-tube electrometer. The counting rate for ordinary rock sources is from 0.3 to 4.0 alphas per mg per hour. The large effective area (150 cm²) of the ionization chamber allows up to 200 mg of sample to be used. The helium content (determined by the direct-fusion method[2]) divided by the rate of production of helium gives a direct measure of the geologic age that is independent of the decay constants of the radioactive series. Age determinations on several rocks have been made.

[1] G. D. Finney and R. D. Evans, Phys. Rev. **48**, 503–511 (1935).
[2] R. D. Evans, R. S. I. **6**, 99–112 (1935); R. D. Evans and C. Goodman, Phys. Rev. **51**, 595 (1937).

17. Focusing X-Ray Ionization Spectrometer. S. T. STEPHENSON, *State College of Washington.*—The ionization method is here applied to the curved crystal x-ray spectrograph. Focusing x-ray spectrographs of the type developed by Cauchois and others are being used in several laboratories. The general method involves transmission through a curved crystal and registration on a photographic plate placed on the focal circle. In the present investigation, an optical spectroscope is used as a base. The curved crystal holder, having a radius of curvature of 2 feet, is mounted in place of the collimator 1 foot from the central axis of the spectroscope. A slit is mounted in place of the telescope at a distance of one foot from the central axis and thus is capable of rotation in the focal circle. The ionization chamber is mounted behind the slit. The slit is moved in small angular steps through the region to be investigated and the ionization currents are recorded for each slit position. Tests on the Mo $K\alpha_1\alpha_2$ doublet indicate better intensity than that obtainable with the usual type Bragg spectrometer for a given resolving power. The mica crystal used has given line widths which are about four times greater than those obtained with a double crystal spectrometer.

18. Liquid *vs.* Vapor Temperature. EDWARD M. LITTLE, *Montana State University.*—It is usually assumed that when a liquid and vapor are in equilibrium, their temperatures are the same. This paper proposes to show they are not, if temperature is defined as proportional to the average kinetic energy of translation of the molecules, and not by its effect on a thermometer. There are two opposing things

that tend to give the vapor a different temperature: (1) only the faster molecules can penetrate the meniscus, (2) the surface tension forces reduce the kinetic energy of the molecules vaporizing. These raise and lower the temperature respectively. Assuming a Maxwell distribution of velocities in the liquid and a definite potential difference (or molecular work function) between vapor and liquid, the average kinetic energy of the molecules before and after they leave the liquid can be found. A thermometer will never detect this temperature difference; it must be tested by methods (such as spectroscopic) which do not change the molecular energy appreciably. We show that vapor temperature may be 2/3 the liquid temperature.

19. Electron Asymmetry in the Atoms of Zinc Crystals. G. E. M. JAUNCEY AND E. M. MCNATT, *Washington University.*—The diffuse scattering of x-rays from single crystals of zinc at room temperature has been measured at scattering angles ϕ from 15° to 40° and at orientation angles ψ of 14° and 90°. Using our S values and Miller's F values obtained from the reflections of Mo $K\alpha$ x-rays from powdered zinc, we have been able to calculate the true atomic structure factors (from which the effect of thermal vibration has been removed) for orientation angles $\psi=0°$ and $\psi=90°$. When these f values are plotted against $(\sin\phi/2)/\lambda$, two curves—one for $\psi=0°$ and the other for $\psi=90°$—are obtained. The two curves coincide for values of $(\sin\phi/2)/\lambda$ greater than 0.6, with the fork occurring at this place. The difference of the f values is due to electron asymmetry and gives a measure of this asymmetry. The method of diffuse scattering is unusually powerful in this type of problem.

20. The Use of Twin Sources in Experimental Studies of Thermal Neutrons. G. J. THIESSEN AND E. L. HARRINGTON, *University of Saskatchewan.*—If using but a single radon-beryllium source in the central portion of a body of a hydrogenous material, the plan generally followed, one obtains thermal neutron radiation which varies greatly with the distance from the source. This variation makes it difficult to obtain consistent results whatever the detector used. By employing two such sources placed at the optimum distance apart the authors were able to obtain a region of nearly uniform neutronic radiation, the scheme being analogous to that employed in magnetism to obtain a uniform field through the use of Helmholtz coils. Certain experimental studies, using such a region, have been carried out.

21. A Deuteron Source for Nuclear Research. N. E. BRADBURY AND F. BLOCH.—A modification of the usual low voltage arc in deuterium for the production of deuterons is described which appears to have several advantages over the ordinary types. A relatively high deuteron yield is obtained by using pressures of 0.1 mm to 0.4 mm and an auxiliary cathode as described by Luhr and Lamar. To avoid excessive gas consumption and high pumping speeds, an electrostatic focusing system is incorporated in the arc which permits currents of the order of 0.4 ma to pass through an aperture of 1 mm². The arc chamber is water cooled. The arc is easy to start and stable in operation with a gas consumption of 20 cc/hour under ordinary operating conditions. The total power required for operation of the arc is 250 watts which may be supplied at high potential by means of an insulation transformer.

22. The Effect of Containing Tubes on Ultrasonic Velocities in Benzene. W. H. HULSWIT, JR. AND B. J. SPENCE, *Northwestern University.*—Field and Boyle[1] have shown that for frequencies up to 110 kilocycles the measured velocity of sound in a liquid is a function of the frequency. A radial resonance phenomenon associated with the containing tube was shown by Field[2] to furnish an explanation.

In carrying this investigation to higher frequencies two interferometers, having tubes of 1 inch and 1½ inches, were employed at frequencies from 400 to 700 kilocycles using toluene free C.P. benzene as the liquid. At these frequencies it was found that the measured velocity of sound was not a constant but passed through definite maxima and minima, with a maximum variation of about 5 percent. The change was cyclic and occurred about every 80 kilocycles with the 1-inch tube and every 50 kilocycles with the 1½-inch tube. The locations of the maxima are in agreement with Field's theory if an "effective" diameter somewhat smaller than the real diameter of the tube is employed in calculations.

[1] Field and Boyle, Can. J. Research **6**, 192 (1932).
[2] Field, Can. J. Research **5**, 131 (1931).

23. A Multiple Camera for Lightning Studies. E. J. WORKMAN AND R. E. HOLZER, *University of New Mexico.*—An improved lightning camera designed to obtain relatively complete photographic records of the flashes in a single storm will be described and exhibited. The instrument consists of three rotating drum units and one still camera unit mounted in a single case in such a way as to provide rapid film changing. The camera is driven by a constant speed six volt motor and film changing on all units is accomplished in about two seconds while the drums are in motion. Standard moving picture film is placed in magazines within the rotating drums in such a way that thirty to forty pictures may be taken without reloading the magazines. The only manual operation necessary in taking a picture and preparing for the next picture consists of closing a single electrical contact. Two of the drums rotate about vertical axes and one about a horizontal axis. A simple and very satisfactory method has been developed for analyzing complicated discharges through photographs taken in this way.

24. Fringe Systems Produced by Passage of X-Rays Through Plane Slits. A. E. SMICK AND PAUL KIRKPATRICK, *Stanford University.*—X-rays emerging from a plane slit system such as that formed by parallel glass plates in contact diverge in a fringe system having maxima and minima parallel to the slit plane, as in ordinary slit diffraction. The principal features of these patterns can not be explained by diffraction, however, as the width and separation of fringes is too great. The role of multiple total reflection has been investigated and been shown to be chiefly responsible for the observed effects.

25. The Production of Ions in the Auroral Glow. J. KAPLAN, *University of California at Los Angeles.*—Studies of the discharges which produce the auroral afterglow, as well as of the afterglows, are brought together in an attempt to construct a hypothesis regarding the production of ions in these glows. The answer to this problem will have an important bearing on the production and decay of ions in the earth's upper atmosphere. The hypothesis is based chiefly on the observation that at high pressures the $v'=1$ level of the initial electronic state of the Goldstein-Kaplan bands (G-K level) is enhanced at the expense of metastable molecules in the $A^3\Sigma$ state. There exists almost an exact two to one ratio between the energies of these states. Collisions between two molecules in the G-K state, or between one in the G-K and one in the $A^3\Sigma$ state, involve enough energy to yield N_2^+. Such properties of the glow as the violent enhancement of the green members of the first-positive bands and the intensity distribution in the second-positive system are ascribed to collisions between molecule-ions and metastable molecules or normal molecules. This hypothesis leaves us with many predictions regarding the variation of intensity in auroral spectra with pressure as well as predictions of hitherto unobserved bands in auroral afterglows. These will be studied.

26. A New Charcoal Trap for Oil Vapors. PAUL A. ANDERSON, *State College of Washington.*—The limitation of preliminary outgassing temperatures to 500°C in glass-contained charcoal traps heated by external furnaces and the non-uniform heating with rapid element deterioration characteristic of immersion heating with wire-wound units have been avoided by making the charcoal mass act as its own heating unit in the manner of the granular carbon resistance furnace. The charcoal is packed in a reentry tube between graphite rods which cap molybdenum current lead wires welded to a conventional tungsten-in-Pyrex press. A charcoal column 2.5 cm in diameter and 10 cm long is brought to incandescence (800–900°C) in a few minutes by currents which do not overload 60 mil tungsten leads. The prolonged preliminary outgassing periods required in conventional traps for adequate clean-up are reduced from days to hours and the vacuum attainable with the charcoal hot is improved markedly.

AUTHOR INDEX TO ABSTRACTS OF DENVER MEETING

Altar, William—see Condon, E. U.
Anderson, Paul A.—No. 6
—— No. 26

Ballard, Stanley S.—No. 7
Bloch, F.—see Bradbury, N. E.
Bowls, Woolford E.—No. 5
Bradbury, N. E. and F. Bloch—No. 21
Broxon, James W.—see Merideth, George T.
Broxon, James W. and George T. Merideth—No. 1

Cohn, Byron E. and S. C. Lind—No. 10
Condon, E. U., William Altar and Henry Eyring—No. 9

Evans, Robley D. and Clark Goodman—No. 16
Eyring, Henry—see Condon, E. U.

Froman, D. K.—see Stearns, J. C.
—— **and J. C. Stearns**—No. 15

Gardner, M. E.—No. 3
Goodman, Clark—see Evans, Robley D.

Harrington, E. L.—see Thiessen, G. J.
Holzer, R. E.—see Workman, E. J.
Hulswit, W. H., Jr. and B. J. Spence—No. 22

Jauncey, G. E. M. and E. M. McNatt—No. 19
Johnson, Thomas H.—No. 14

Kaplan, Joseph—No. 12
—— No. 25
Kirkpatrick, Paul—see Smick, A. E.

Lagsdin, J. B.—see Wyckoff, Ralph W. G.
Laporte, Otto—see Mack, J. E.
Lind, S. C.—see Cohn, Byron E.
Little, Edward M.—No. 18

Mack, J. E. and Otto Laporte—No. 11
McNatt, E. M.—see Jauncey, G. M.
Merideth, George T.—see Broxon, James W.
—— **and James W. Broxon**—No. 2

Smick, A. E. and Paul Kirkpatrick—No. 24
Spence, B. J.—see Hulswit, W. H., Jr.
Starr, Merle A.—No. 4
Stearns, J. C.—see Froman, Darol K.
—— **and D. K. Froman**—No. 13
Stephenson, S. T.—No. 17

Thiessen, G. J. and E. L. Harrington—No. 20

Workman, E. J. and R. E. Holzer—No. 23
Wyckoff, Ralph W. G. and J. B. Lagsdin—No. 8